GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY
CARL MURCHISON

VOLUME 25
1942

Published quarterly by The Journal Press
Provincetown, Massachusetts, U. S. A.

TABLE OF CONTENTS

No. 1

No. 2

The Journal of Psychology

EDITED BY

Carl Murchison

The Journal of Psychology was founded at the request of many psychologists throughout the world for the purpose of establishing a medium in which publications shall be immediate and in which other publication experiments may be carried out.

(1). The subscription price will be $7 per volume.

(2). The format of each volume will be such that each subscriber may tear down the assembled volume into its component articles.

(3). Except through the fault of the author, the lag in publication should not be more than three or four weeks.

(4). The author will receive 200 reprints of his paper, and the subscriber will receive the accumulated publications bound into half volume units. The second half of each volume will contain the title page and table of contents for that volume. Each half volume unit will be mailed to the subscribers at its stated date.

(5). Each author will be charged the actual manufacturing costs of any zinc or half-tone cuts.

(6). The publication charges will be as follows:

4	page	signature	$12
8	"	"	24
12	"	"	36
16	"	"	48
20	"	"	60
24	"	"	72
28	"	"	84
32	"	"	96
36	"	"	108
40	"	"	120

These charges are tentative and will be adjusted as justified by the subscription list (especially by the effect of the war in China and in Europe).

(7). The accumulated units will be mailed to subscribers quarterly: January, April, July, and October. The units may vary in size, to accommodate publication needs.

Two volumes annually. Per annum $14.00; per volume $7.00; per half volume unit $4.00. Complete sets at $7.00 per volume plus transportation.

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

$7.00 per volume
Single numbers $4.00

QUARTERLY
Two volumes per year

February, 1942
Volume 25, First Half

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

FEBRUARY, 1942

(Manuscripts are printed in the order of final acceptance)

Provincetown, Massachusetts
Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under the act of March 3, 1879
Reentered as second-class matter April 15, 1937, at the post-office at Provincetown, Mass., under Act of March 3, 1879.

Published as a separate and in *Genetic Psychology Monographs*, 1942, 25, 3-99.

A STUDY OF FACTORS DETERMINING FAMILY SIZE IN A SELECTED PROFESSIONAL GROUP*

JOHN C. FLANAGAN

Coöperative Test Service of the American Council on Education

*Received in the Editorial Office on August 10, 1941, and published at Provincetown, Massachusetts.

ACKNOWLEDGMENTS

The investigation reported in the following pages was undertaken at the request of the Pioneer Fund. Acknowledgment is especially due to Mr. Frederick Osborn, Secretary of the Board of Directors, for assistance throughout the study.

The collection of the data from the personnel of the United States Army Air Corps was greatly facilitated through the coöperation of the Secretary of War and of General H. H. Arnold. Acknowledgment is also due Major Loyd E. Griffis, a member of the Medical Corps attached to the Army Air Corps, who visited the various fields and explained the study to the officers personally while on leave from active duty.

The valuable assistance of the members of the Advisory Committee who assisted in the preparation of the forms and with the planning and preliminary tryout of the procedures to be used in obtaining the data is gratefully acknowledged. This committee included the following persons: Dr. Elmer D. West of Stoneleigh College, Dr. E. Lowell Kelly of Purdue University, Dr. Frank Shuttleworth of the College of the City of New York, Dr. Neal Miller of the Institute of Human Relations of Yale University, Mr. Charles R. Langmuir of the Carnegie Foundation, and Major Loyd E. Griffis of the U. S. Army Medical Corps. The members of this committee also very generously read and criticized the preliminary draft of this report.

Acknowledgment is also gratefully made to Dr. Margaret E. Erb and Miss Leone E. Chesire, who supervised the tabulation and analysis of the data, and to Mr. Robin M. W. Travers and Dr. Elmer D. West who prepared the preliminary draft of the present report.

JOHN C. FLANAGAN

5616 Oak Place
Bethesda, Maryland

I. INTRODUCTION

The rapid decline in this century of the birth rate in the United States and the nations of Europe has led to a number of investigations in this field. Most of these studies have been concerned with the differences between the birth rates of social, geographical, or economic groups. The present study represents a preliminary effort to investigate certain social, economic, and psychological factors affecting size of family by means of an intensive study of one relatively small professional group.

This study is limited to officers of the United States Army Air Corps and their wives. The group is therefore quite homogeneous, distinctly above the average of the general population in measurable mental and physical traits, and the cultural and economic background of the individuals in it is fairly similar. It must, however, be borne in mind that even in the relatively normal period in which this preliminary study was made (i.e., before the present world emergency) this group may have been influenced by considerations *not equally operative in the general population.*

The immediate purpose of the preliminary study conducted in the spring of 1938 was to obtain data on a select group, the officers of the United States Army Air Corps, with particular reference to: (*a*) how many children they have, i.e., the extent to which they are reproducing themselves; (*b*) how many children they would like to have, or what they consider to be the ideal number of children per family; (*c*) basic psychological, biological, social, economic, and other factors which account for the number of children they have and the number they plan to have and the discrepancy between these and their ideal, i.e., why they would like a certain number of offspring, whether they have this number, and if not, what are the reasons; (*d*) practical measures, based upon elements which appear to affect the family size, which would lead to an increase in the number of offspring.

That attitudes and personality traits are of marked importance in determining family size is evidenced by a consideration of some of the factors involved. Many married couples, when asked why they have no children or why they have no more children, answer by saying that they cannot afford the added expense. The answer is, at least superficially, an economic one. It is frequently, however, not an economic factor in terms of dollars of income, but only so in

terms of relative economic status. This immediately leads one to the problem of standards of living for various occupational groups; to attitudes toward essential expenditures, and to other attitudes and personality characteristics. Within a given occupational field or economic level, varying birth rates must be explained in other than economic terms. Clearly, health and sterility will be involved for some, but after these are eliminated differences will still be found. In making a choice between a baby and a new car, what factors determine the decision? The car may be chosen for personal pleasure, to keep up with the class, to give more freedom for social purposes, to escape responsibility, because of fear of childbirth or of defective children, or for a number of additional reasons. These will rarely be expressed. They may not, in fact, even be recognized. The financial situation plays a part, but the *attitude* toward the income and other factors may be equally or more important.

As mentioned above one of the basic purposes of the preliminary study was to obtain information concerning the practical measures which might be employed to assist those individuals in a group such as this who so desired to have larger families. As a result of recommendations based on the findings of the preliminary study educational scholarships were made available to children born to regular officers in the Army Air Corps during the calendar year 1940 provided there were already at least three children in the family.

The text of the leaflet announcing the availability of these scholarships is given in Appendix *A*.

In the spring of 1939 prior to the time these scholarships were offered the Board of Trustees of the Pioneer Fund requested that a statement concerning the expected number of children who would qualify for such Educational Scholarships if offered be presented to them by the director of the field study. In response to this request the following statement was submitted to them in April, 1939.

> The mean number of children who would have qualified in each of the past 10 or 11 years is 4.1. In each of these years there have been approximately 60 officers with three or more living children and the youngest 5 years of age or less. The expected fluctuation from year to year, using the standard statistical formula for the chance fluctuation of an event with a probability of 4 in 60, is found to be 1.9 (standard deviation). That this theoretical value is in fairly good agreement with the facts is shown by the calculated standard deviation of the

numbers of children who would have qualified in each of the past 11 years. The calculated standard deviation is 2.5. It is therefore seen that there is less than one chance in a thousand that as many as 12 children would be born in any particular year without the intervention of some fundamental change in the conditions.

A very careful analysis of the replies given by the officers and their wives in the recent survey indicates that of the 33 officers included in the sample who had at least three living children, about four are quite certain to accept the offer of the scholarship, six appear to be quite favorably inclined and should be considered as at least strong possibilities, and four others although not very likely to have another child should at least be considered as possibilities. Allowing for the fact that 18 of the 33 have a child under 6 years of age, and using the figures obtained from the Personnel Division indicating that about 60 officers with three children or more have a child under 6, we estimate that these officers and the other 80 officers with at least three children will be very likely to have at least 12 children born during the calendar year 1940; that it is probable that they will have as many as 20 or 30 children during this period; and that there is a remote possibility that as many as 50 children might be born. This latter possibility seems to be extremely remote however and the chances of its happening are certainly less than one in a thousand.[3]

There were 12 children born in 1940 who qualified for scholarships under these conditions. The April, 1939, statement predicted that at least 12 children would be born during the calendar year 1940 if the scholarships were offered and also pointed out that there was less than one chance in a thousand that this would occur unless the scholarships were offered.

There appear to be two principal reasons why this figure was somewhat below certain of the higher estimates given in 1939 on the basis of the analytical study which had been made. First, the present world war, which introduced a larger factor of uncertainty into the lives of these officers than would ordinarily have been the case; and second, the fact reported by individuals in the Medical Corps, that a number of officers and wives who indicated that they

[3]Flanagan, John C. Statement concerning expected number of children who will qualify for the Educational Scholarships which it is proposed that the Pioneer Fund offer as presented to the Board of Trustees of that organization at its meeting April 10, 1939.

would have another child if scholarships were made available found that they were physiologically incapable of bearing children within the time limits given.

It seems clear from these findings that the financial factor is not a negligible one in determining family size but similarly it is not as predominant as a superficial investigation might imply.

The present study does not pretend to untangle the complex matrix which includes the factors responsible for the determination of family size. It is being reported mainly in the hope that the procedures developed and some of the preliminary indications as to types of factors which may be of importance will be found useful by later investigators. The reader who is interested in a systematic review of the literature or a general discussion of this problem is referred to the excellent recent book by Osborn (7) and an earlier book by Lorimer and Osborn (5).

II. GENERAL PROCEDURES AND DESCRIPTION OF SAMPLE

A. General Plan of the Study

As previously noted the plan of this study differed rather fundamentally from those of previous investigations in this field. Earlier studies have been concerned with one or two or at the most only a few items of information about each individual but have included thousands of individuals—frequently whole populations of regions. In the present study over two thousand separate items of information were obtained concerning each family but only a few hundred families were included.

This study differs from most previous studies not only in the amount but also in the type of information secured from each individual. While most studies have been confined to the collection, tabulation, and analysis of objective data such as age, race, education, religion, occupation, income, size of community, and age at marriage, the present investigation also includes many subjective and psychological factors such as attitudes, interests, preferences, and values.

The relevance of such psychological data to an investigation of family size is obvious. The principal question concerns the validity and accuracy of the data and of the interpretations based on an analysis of them. Significant improvements have been made in recent years in the techniques for obtaining and analyzing such data. Because of these developments it seems likely that there will be a number of important studies in this field in the near future.

To assist in formulating the specific procedures and forms to be used in the study an advisory committee was selected. The committee included Dr. Elmer D. West of Stoneleigh College, Dr. E. Lowell Kelly of Purdue University, Dr. Frank Shuttleworth of the College of the City of New York, Dr. Neal Miller of the Institute of Human Relations of Yale University, Mr. Charles R. Langmuir of the Carnegie Foundation, and Captain Loyd E. Griffis of the U. S. Army Medical Corps.

The committee discussed a number of possible methods of collecting the necessary data. The procedures most seriously considered were the interview method and the questionnaire method. It was decided that the questionnaire method would be tried first because of the excellent results Dr. Kelly reported that he had been getting with a modified questionnaire procedure in a somewhat similar situation,

because the questionnaire had been reported as more valid in obtaining personal data than the interview (3), and finally because it would require much more time to secure good results by means of the interview procedure than for the questionnaire method, both for the investigators and the officers and their wives.

The committee prepared a preliminary form of the questionnaire for which a few of the items were taken from forms which have been used in studies of marriage (4). Certain published tests (1, 2, 8, 9) were also selected for use with the preliminary groups. In preparing the materials principal emphasis was placed on making the questions as simple and easily answered as possible. Most questions were to be answered by a single word or number and many of them required only the placing of a check mark after the appropriate choice of several given.

B. Preliminary Testing of Materials and Procedures

The general plan and the preliminary set of materials were tested with three small groups: a group of six officers and their wives at Bolling Field, 24 officers and the wives of about two-thirds of them at Mitchel Field, and a group of 20 professors and their wives at the University of Connecticut. The results obtained from these groups were analyzed to obtain suggestions for the improvement of procedures and specific items used. The committee had been somewhat uncertain concerning the extent to which some types of questions would be answered. There were also directions for indicating answers to certain groups of questions, the clarity of which needed testing.

For example, preliminary tests showed that accurate information concerning the sizes of families of uncles and aunts, that is, the number of cousins either ever-born or now living cannot be obtained from many individuals. Similarly it was found that in spite of much careful revision of certain directions by the committee, they were still misinterpreted in certain ways by some of the group. Other information was obtained from the preliminary groups such as that everyone responded to certain questions with practically the same answer. Since this study is mainly concerned with establishing differences, such items were revised or eliminated. It was found that the answers to certain other questions were difficult to interpret and did not contribute to the study as it was expected they would. Such items were also revised or eliminated.

In general it may be said that the committee regarded this preliminary testing as almost an indispensable step for a study of this type. Certainly the questionnaire called the Family Report Form was very significantly improved as a result of this preliminary work, and it was possible to proceed in the planning of the main study with confidence that the procedures were practical and could be expected to yield useful results.

C. Collection of Data

The general procedure adopted was to arrange for the personal presentation of the general plan of the study to officers at the various posts and request their coöperation and that of their wives in providing the necessary information. Since it did not seem practical to visit all of the 40 or 50 posts at which the officers in the Army Air Corps were stationed, it was planned that a sample be obtained from the larger fields. An indication of the way in which the officers were distributed is provided by the following figures giving the number of Air Corps officers stationed at the various fields in March, 1937: Randolph Field, Texas, 154; Langley Field, Virginia, 122; Maxwell Field, Alabama, 109; Barksdale Field, Louisiana, 84; March Field, California, 83; Wright Field, Ohio, 70; Chanute Field, Illinois, 68; Kelly Field, Texas, 67; Hamilton Field, California, 65; and Mitchel Field, New York, 58.

It is seen that these 10 fields account for a little less than 900 or about 60 per cent of the total group of 1472 regular officers who were in the Army Air Corps at that time. *The remaining 600 were divided into small groups with about 10 or 15 officers in each group which were scattered over the United States and its territorial possessions.*

Through the coöperation of various individuals in the War Department and especially General H. H. Arnold, then Assistant Chief of the Air Corps, it was arranged that Captain Loyd E. Griffis, a flight surgeon on leave from the U. S. Army Medical Corps, should visit the fields selected and explain the study in some detail to the Commanding Officer. The Commanding Officer then made arrangements, insofar as possible, for Captain Griffis to present the plan of the study to the officers at that particular post. Although the participation of the officers was on an entirely voluntary basis, excellent coöperation was obtained. Of the group of approximately 600 officers to whom the study was presented 427 officers and 326 of their wives filled in the forms.

This response appears to be especially good when it is noted that the complete set of forms and tests usually required more than three hours to fill out. A complete summary of the forms used and the responses obtained is given in Appendix *B*.

It is difficult to decide which factors were the most important in enlisting the coöperation of the group, but those of particular value included the intrinsic interest of the problem, the precautions to ensure anonymity, the offer of a report on the personality and interest blanks, the official approval which had been given to the project, and the possibility of obtaining a vocational interest pattern on the Strong *Vocational Interest Blank* which might be of assistance in the selection of future officers.

The officers were asked to enlist the coöperation of their wives. Each officer who was willing to participate signed for a set of forms for himself and for his wife with the instructions to return them sealed to the commanding officer when completed, or to mail them, in the envelope provided, direct to the office where they were to be analyzed.

The officers were told that on all but the purely factual items independent opinions from husband and wife were desired. Although it is recognized that most of the couples must have previously discussed many of the items, the results indicate that the opinions reported were nearly always given independently.

D. Representativeness of the Sample Obtained

If generalizations are to be made concerning the officers in the Army Air Corps on the basis of the data obtained from this sample, it is important to examine the evidence concerning the representativeness of the sample. In the first place the probable effect of the procedures used in collecting the data will be discussed. To what extent may the total group of officers stationed at the particular posts visited be considered as representative of all officers in the Army Air Corps? Since all of the large fields as well as a few of the small fields were visited, the only possible selective factors introduced in this way would appear to be those associated with size of post. An important factor here is the general policy of the Army to have officers change posts every four years. In recent practice most officers in the Air Corps are transferred every two years. This results in quite a general shuffling and would seem to increase very greatly the chances that the sample obtained should be representative of the Air Corps.

The next point concerns the extent to which the group filling out the forms is typical of the entire group of officers at a particular field. It appears likely that the group actually seen at a given field would tend to be somewhat more representative of the younger officers than of the older ones. The higher ranking officers would be somewhat more difficult to bring together and presumably also have less free time to devote to such a project than the younger officers. However, the variety of appeals used in gaining the coöperation of the officers as mentioned above should serve to prevent the responding group from being seriously biased.

Some actual evidence concerning the representativeness of the sample in certain respects is possible by comparing the sample with the total Air Corps on those characteristics for which data on the total group is available in the personnel files of the Air Corps.

Table 1 shows a comparison of the ranks of the officers in the sample and in the entire Army Air Corps.

TABLE 1

COMPARATIVE DATA CONCERNING RANKS OF THE OFFICERS

Rank	Number		Per Cent	
	Army Air Corps 1936	Sample 1938	Army Air Corps 1936	Sample 1938
Colonel	44	7	3	2
Lt. Colonel	82	10	6	2
Major	332	61	24	14
Captain	304	111	22	26
1st Lieutenant	508	180	36	42
2nd Lieutenant	127	55	9	13
Unknown and other	0	3	0	1
Total	1397	427	100	100

With respect to rank, it is to be noted that the sample includes a larger proportion of the officers of the lower ranks than did the total Air Corps in 1936. This is to be expected since the older officers cannot be expected to be as interested or as able to find time to coöperate as the younger group. However, this does not appear to be very serious in relation to the general purposes of the investigation. It might, however, tend to introduce a bias into the older groups which would interfere with comparisons of the responses of groups of young and old officers. This should be kept in mind in interpreting the data presented in later chapters.

TABLE 2

COMPARATIVE DATA CONCERNING MARITAL STATUS BY AGE CLASSIFICATIONS

	Number						Per Cent					
	Married or divorced		Single		Total		Married or divorced		Single		Total	
Age	Army Air Corps 1936	Sample 1938	Army Air Corps 1936	Sample 1938	Army Air Corps 1936	Sample 1938	Army Air Corps 1936	Sample 1938	Army Air Corps 1936	Sample 1938	Army Air Corps 1936	Sample 1938
60-64	3	0	0	0	3	0	0	0	0	0	0	0
55-59	12	2	1	0	13	2	1	1	1	0	1	0
50-54	36	11	4	2	40	13	3	3	3	5	3	3
45-49	141	31	3	0	144	31	11	8	2	0	10	7
40-44	312	53	12	1	324	54	25	14	8	3	23	13
35-39	170	81	11	6	181	87	14	21	7	16	13	20
30-34	338	148	38	10	376	158	27	38	26	26	27	37
25-29	219	59	72	18	291	77	18	15	49	47	21	18
20-24	18	4	7	1	25	5	1	1	5	3	2	1
Total	1249	389	148	38	1397	427	100	100	100	100	100	100

In Table 2 are presented comparative data concerning age and marital status. It is seen that the sample contains about 10 per cent more officers' ages 30-34 and about 10 per cent fewer officers' ages 40-44 than does the entire Army Air Corps. Otherwise the groups are quite similar with respect to these factors.

The only other characteristic on which data were available for the whole of the Air Corps was family size. As will be seen in Table 3 the sample appears fairly similar to the total Air Corps in this respect.

TABLE 3

COMPARATIVE DATA CONCERNING FAMILY SIZE

	Number		Per Cent	
	Army Air Corps 1936	Sample 1938	Army Air Corps 1936	Sample 1938
		Single		
	148	38	11	9
		Married or divorced		
0 children	435	117*	31	27
1 child	397	125	28	29
2 children	296	101	21	24
3 children	83	28	6	7
4 children	29	12	2	3
5 children	5	2	0	0
6 children	3	3	0	1
7 children	1	0	0	0
10 children	0	1	0	0
Total	1397	427	100	100

*This group includes 38 officers unable to have children because of sterility, six officers who are both childless and divorced, 12 who are expecting their first child, and 61 others who presumably could have or could have had children.

The sample has a slightly greater number of children per officer. Some of this apparent difference is due to the fact that the sample includes cases of stillbirths whereas the figures given for the whole of the Air Corps in 1936 do not.

Table 4 gives the comparative proportions of West Point graduates in the total group of Air Corps officers, as of January 1, 1937, and in the sample of officers obtained in this study. It is seen that the proportion of graduates of the United States Military Academy is somewhat greater in the sample than in the Air Corps as a whole.

Although not perfectly representative of the total Air Corps, this sample of officers includes officers representing all of the broad

TABLE 4

COMPARATIVE PROPORTIONS OF GRADUATES OF THE UNITED STATES MILITARY ACADEMY AT WEST POINT IN THE TOTAL GROUP OF AIR CORPS OFFICERS, AS OF JANUARY 1, 1937 AND IN THE SAMPLE OBTAINED IN THE SPRING OF 1938

	United States Military Academy	Other
Total Air Corps (1937)	31%	69%
Sample (1938)	41%	59%

categories of rank, age, and family status present in the total group. There is no evidence that the method of sampling or the procedures used in collecting the data have introduced any serious bias which might invalidate the general findings of this study.

E. Description of the Group

As was seen in the previous section a large number of the officers in the Air Corps are graduates of the United States Military Academy at West Point. Most of the other officers are graduates of colleges. In addition to such academic training these officers are normally graduates of the Army's Flying School. They are not allowed to marry until they have completed their flying school course. For this reason the average age at marriage is somewhat greater for this group than for most groups. However once they have received their commissions in the Army Air Corps, they are established with a regular salary and the expectation of regular increases and a relatively adequate retirement income.

As will be seen in Table 5, most of the officers get married fairly soon after their training is completed. This table also shows that most of the officers are between 25 and 50 years of age. As indicated by the median years of service of the various age groups they usually enter the service at about 22 or 23 years of age. By the time they are 30 about 90 per cent of them are married and very, very few of them never marry. Rank is almost entirely a matter of length of service. An officer in his forties is usually a major and an officer in his fifties a colonel in the Air Corps. Salary is in accordance with the schedules published in the official *Army Register*. Base pay is determined by rank and length of service. A living allowance is provided, the amount of which is determined by rank, length of service, and whether the officer has dependents. At many Army posts housing facilities are provided by the Army for the

TABLE 5

DATA CONCERNING PER CENT MARRIED, RANK, YEARS OF SERVICE, SALARY, AND AVERAGE NUMBER OF CHILDREN FOR OFFICERS IN THE ARMY AIR CORPS CLASSIFIED BY AGE GROUPS

					Monthly rate of pay including flying pay and living allowance				
Age group	No. of cases	Per cent married	Rank (Mode)	Years of service (Median)	With dependents (Mode)	Without dependents (Mode)	No. of children per family with children	No. of children per married officer	No. of children per officer
60-64	5	100	Colonel	40	906	848	2.00	1.33	1.33
55-59	13	92	Colonel	35	906	848	2.86	1.67	1.54
50-54	40	90	Colonel	29	906	848	1.89	1.42	1.28
45-49	144	98	Major	22	641.50	565.50	1.92	1.29	1.26
40-44	324	96	Major	19	641.50	565.50	1.96	1.52	1.46
35-39	181	94	Captain	14	476	438	1.82	1.32	1.24
30-34	376	90	1st Lt.	10	383.50	345.50	1.45	0.96	0.86
25-29	291	75	1st Lt.	5	358.50	320.50	1.27	0.53	0.40
20-24	25	72	2nd Lt.	2	245.50	245.50	1.00	0.11	0.08
Total	1,397	89					1.72	1.12	1.00

amounts allowed the officers for such purposes. The Air Corps officers receive 50 per cent additional base pay when on flying duty. In order to qualify for this additional amount, it is necessary that they be ordered to flying status and fly at least four hours every month. Most air corps officers are on flying status and do qualify.

It will be noted that the salary schedule compares favorably with those of successful professional groups. There is no insurance provided free by the government for Army officers, neither are the officers compelled to carry any; but within 120 days of obtaining their commission they may take out $10,000 of government War Risk Insurance given at a rate which compares very favorably with that provided by private companies for civilians. In practice most of the Air Corps officers carry substantial amounts of life insurance, underwritten by private companies. Some private insurance companies will not sell life policies to men engaged in aviation owing to the difficulties of estimating the risk. Those which do issue such policies require an extra premium of about $10.00 per $1,000 of insurance per year.

The government does provide gratuity pay and pensions. In general, gratuity on death amounts to six months' base pay in a lump sum, including flying pay and pay for qualifications in the use of arms, payable to widows and children under 21 years of age and unmarried, or otherwise to designated dependent relatives. Pensions amount to $22.00 for a widow under 50 years of age; $26.00 for a widow 50 to 60 years of age; and $30.00 for a widow over 65 years of age. A widow with one child gets $7.00 additional pay to 10 years, increase to $11.00 from age 10, with $6.00 for each additional child up to 10 years. When there is no widow, but one child, $15.00; no widow but two children, $24.00; no widow but three children, $34.00, equally divided. No monthly rate exceeds $56.00 regardless of the number of dependents.

Medical attention is provided by the Army for all military personnel and for civilian dependents (wife and children) where possible. After 30 years' service officers can ask for retirement. Such requests are ordinarily granted. In practice quite a number of officers in the Air Corps are retired earlier because of the high physical requirements of flying.

The average number of children per officer for men over 40 years of age as found from Table 5 is 1.39 which is lower than that of most professional groups. This number will be reduced still

TABLE 6

A Comparison of Certain Age Groups of Army Air Corps Officers with Various Other Groups with Respect to Per Cent Ever Married and Size of Family

Group	Year in which data were obtained	Number of cases	Approximate age	Per cent ever married	Per cent childless marriages	Mean children per marriage
Total Army Air Corps	1936	504	40-55	96.1	26	1.46
Swarthmore classes	1930	303	40-56	88.4	21	2.15
Harvard classes	1924-26	1,899	47	81.4	25	1.9
Yale classes	1918-22	692	47	82.4	25	1.9

TABLE 7

A Comparison of Reports of Ideal Number of Children Obtained from Various Samples

Source	Group	Sample	Mean	Ideal size in per cent 0	1	2	3	4	5	6	7
Present Study	Wives	310	3.24	0	1	22	37	37	3	1	0
	Officers	417	3.24	0	0	19	46	29	5	1	0
Ladies' Home Journal March, 1938	Women Readers	"Representative" Sample of Women Readers	3.33	0	0	25	32	34	4	4	1
	All Women	"Representative" Sample of U. S.	3.41	0	1	26	25	36	5	5	2
American Institute of Public Opinion December, 1936	American Men and Women	"Representative" Sample	3.17	0	2	32	32	22	7	3	2
	Relief Workers	"Representative"	3.53	0	1	24	29	25	12	6	3
Youth Tell Their Story H. M. Bell, 1938	5,000 Young Men Aged 16-24	"Representative" Sample of the Youth of Maryland	2.40	10	8	44	20	12		←6→	
	5,000 Young Women Aged 16-24		2.48	8	8	44	20	12		←7→	

further before the children are old enough to play their part in the reproductive cycle.

The birth rate of this group is considerably lower than those which have been reported for most other groups. In a survey of Columbus and Syracuse Notestein and Kiser reported (6) that the sizes of completed families of professional people were 2.40 and 1.88 respectively. For these same cities the completed families of people in business were 2.56 and 2.18, for skilled workers 2.71 and 2.29, and for unskilled workers 3.52 and 4.64.

A comparison of the reproductive rate of this group with that of the United States as a whole and with other groups may be made by calculating the net reproductive rate used by Lorimer and Osborn (5). The value found for the Air Corps by this method is 0.54 whereas the rate for the total population is reported as 1.10. The Air Corps' net reproductive rate is also below that of the other professional groups which collectively have a value of 0.76. According to Lorimer and Osborn there can be no doubt from this that the officers of the Air Corps are a group which will not have proportional representation in the next generation. It should be mentioned that the net reproductive rate takes into account age and differential mortality.

Another interesting comparison may be made between the percentage ever married and the number of children per marriage among Air Corps officers and among various groups of college graduates. The figures extracted from Lorimer and Osborn are given in Table 6 with those groups which are comparable with respect to age.

It has been mentioned that Air Corps officers usually move once every two to four years. Other factors which should be noted in connection with the low birth rate shown in Tables 5 and 6 are the lack of educational facilities in the vicinity of some of the Army posts and the hazards of flying. One other point of some importance which should be kept in mind in interpreting the materials in the following chapters is that unpublished data in the files of the War Department indicate that the average size of family of Air Corps officers is about the same as for officers in other branches of the Army.

III. THE CONCEPT OF IDEAL FAMILY SIZE

The data collected in this study may be classified into two principal categories: first those which refer to factors involved in the formation of a concept of ideal family size, and second those which are concerned with the determiners of actual family size under present circumstances. Information obtained of the first of these two general types is presented in this chapter.

Some previous studies have been concerned with certain physical and environmental factors effecting family size, but practically no investigations have been made of the effect of psychological factors. One of the principal aims of the present study was to determine the size of family which would be desired if all physical and environmental factors were made "perfect." This problem is purely a psychological one dependent on personality traits, interests, attitudes, values, and beliefs and is independent of such factors as health, restriction of activities, financial resources, characteristics of spouse, or the uncertainties of life. This chapter represents an analysis of the psychological factors related to the desire to have a number of children and the following chapter is concerned with the factors which cause that number to be modified in actual practice.

A fairly complete summary of the basic data obtained from the Family Report Form and the other psychological report forms used in this study is provided in Appendix *B*. The reader is referred to that summary for a more complete presentation of the reports on which the discussion in this and the following chapter is based.

A. Numerical Reports of Ideal Family Size

The data about which the discussion of this chapter centers is provided by the answers to the third question in Section D of the Family Report Form: *"What do you think is the ideal number of children for the average American family?"* As reported in Appendix *B* the average of the values indicated by the officers was 3.24 children. The average of the values reported by the wives was found to be exactly the same to two decimal places. It is very significant for interpreting these findings to note that approximately 95 per cent of both officers and wives selected two, three, or four as the "ideal number of children for the average American family." The degree of homogeneity in these answers very seriously reduces the possibility of finding any large and dominating factors causing individual differences in "ideal family size."

To obtain perspective for these values they are shown in Table 7 along with similar data which have been obtained for other groups. In the present study and also in the survey of women made by the *Ladies' Home Journal*, it was found that the average response given by individuals under 30 years of age was 3.0. This is somewhat lower than the average of the older group. It will be seen that most of the values given in Table 7 are in fairly close agreement except those found in the Maryland youth study. The author of the latter study also reports a trend with age in his study. It seems likely that part of this increase with age is due to maturity and other factors accompanying increasing age. It also seems likely that part of it is due to a changing point of view in the oncoming generations. A more detailed study of trends in ideal family size should be of considerable value for predictions of population growth.

Since the concept of ideal family size is found to be fairly similar in rather widely varying groups in this country and shows only a very small amount of variability, it would seem to be to a large degree a social rather than an individual concept. It seems pertinent to examine such data as is available in this study concerning the extent to which this concept is traceable to the family group, the larger social groups, and the characteristics of an individual's personality which are chiefly dependent on physical inheritance rather than environmental acquisition.

B. Factors Associated with Family Background and Training

One of the most obvious points for investigation is the relation of the individual's concept of ideal family size to the number of children in the family of which he was originally a member. A definite correlation is found here. The Pearson product-moment correlation coefficient between reported ideal size of family and the number of brothers and sisters the individual had is .33 for the husband and .23 for the wives. Both of these coefficients are sufficiently large to preclude the possibility of their being considered errors of sampling. A similar analysis shows that the chances are greater than two to one that an officer who reports two or less as the ideal family size comes from a family of three or less children. In the same way it is found that the chances are also more than two to one that an officer who reports four or more as the ideal family size comes from a family of four or more children. All four correlations of

ideal size of family with number of parents' siblings are also positive though somewhat smaller in size. It is seen that an important factor in determining the individual's concept of ideal family size is the size of the family from which the individual comes.

It would seem reasonable to suppose that the concept of ideal family size would be influenced not only by the number of brothers and sisters but also by such other factors connected with the family as happiness of the parents and the individual during childhood. There is a slight tendency for both the officers and the wives who indicate relatively large ideal family sizes also to report a greater than average degree of marital happiness on the part of the parents. On the other hand it is found that in the case of both officers and wives those indicating relatively large ideal family sizes report their childhoods as slightly less happy than do the remainder of the group. Although in neither of these cases is the relationship sufficiently marked to lend a satisfactory degree of confidence to the findings, the fact that it is expected that the relationship would be slight, if present, would seem to justify further investigation of these points. One further point of this general nature was investigated. The individuals were asked whether their parents considered their families to be too large. The reports of both the officers and wives giving four as the ideal family size were practically the same as those giving two as the ideal size. However, since the families referred to in one case averaged four or five children and in the other case only about three, the report that the parents in the two groups considered their families to be of about the right size equally frequently is in itself of significance.

Probably the other factor associated with the home of most significance according to the results of this investigation is religious preference and religious training. In this study it was found that the particular denomination in which membership was held made very little difference in the reported ideal size of family. On the other hand there is a significant relationship between report of extent of early religious training and ideal family size for the wives and a significant relationship between report of present church attendance and ideal family size for the officers. The corresponding relationships for officers and wives respectively are similar but not sufficiently marked to yield a satisfactory degree of confidence when the usual statistical tests are applied.

In Section *G*, which is discussed in detail later in this chapter,

one of the items called for a report of the importance to them personally, as a reason for having children, of the statement: *"Having children is considered a religious obligation."* In response to this question 51 of the 427 officers and 53 of the 320 wives indicated that they considered it of Some, Much, or Great Importance. Most of the officers and wives considered other reasons as of much greater importance.

One other type of information was obtained concerning the relation between attitude toward religion and ideal family size in connection with the Allport-Vernon *Study of Values Scales.* Scores on the Religious Scale of this blank were found to be significantly related to ideal family size as reported by both officers and wives. This scale provides a means of comparing interests in various objects and types of activities in an effort to determine those patterns of values which are dominant in the particular individual. The Religious Scale is intended to represent religion in the broadest sense as an effort to "comprehend the cosmos as a whole." A high score on this scale does not necessarily represent the acceptance of any particular creed or dogma.

Another factor relating to ideal family size which is associated with home training is the general question of perpetuating the family. Twenty-seven per cent of the officers and 22 per cent of the wives indicated that they regarded the carrying on of the family name and social tradition as of "much" or "great" importance to them as a reason for having children. Nine per cent of the officers and 6 per cent of the wives reported that carrying on the professional traditions of the family was considered of "much" or "great" importance by them as a reason for having children. In responding to the remaining question related to this topic, 5 per cent of the officers and 3 per cent of the wives checked "carrying on the family enterprises" as of "much" or "great" importance to them as a reason for having children. It can be concluded from these data that the perpetuation of the family is a factor but not a large one in determining ideal family size.

C. Social Factors Related to Ideal Family Size

The principal evidence concerning the importance of social factors in determining ideal family size is also obtained from the section of the Family Report Form just quoted, Section *G,* which was devoted to *Reasons for Having Children.* Of the various reasons

for having children which might be classified under the general heading of social reasons, *"Good stock ought to reproduce itself"* is checked as of "much" or "great" importance to them by 52 per cent of the officers and 57 per cent of the wives; *"Children are needed to carry on the traditions of the country"* receives a similar response from 35 per cent of the officers and 42 per cent of the wives; *"To perpetuate the social class of which the parents are members"* obtains such a response from 21 per cent of the officers and 17 per cent of the wives; and *"Having children is considered a social obligation"* is similarly marked by 9 per cent of the officers and 10 per cent of the wives. A majority of the group consider it very important that good stock reproduce itself, but do not consider the general social obligation to have children of much importance to them.

Two related reasons which, while not definitely related to social factors appear to be largely influenced by social customs and beliefs, were regarded as very important by this group. Seventy-four per cent of the officers and 83 per cent of the wives reported that the statement, *"A family is not complete without children"* was of "much" or "great" importance to them as a reason for having children, and 33 per cent of the officers and 47 per cent of the wives checked, *"One of the purposes of marriage is to produce children,"* in a similar way.

The predominant reason for having children in this group is that it is the natural and expected thing. One of the traditional responsibilities of being a wife has been to bear and rear children and that continues to be considered a primary duty. This has, of course, been related to the position of the husband as the provider for the partnership. There is evidence that in cases in which the wife has a substantial independent income of her own and therefore is not dependent on her husand for support, she apparently also feels less obligated to have children. This appears, however, to be a relative matter and if the husband has an outside income greater than the independent income of his wife, her status appears to be quite similar to that of the wife without an independent income. In general, it may also be said that a woman who marries a man from a family with a higher socio-economic status than that of her own family is more likely to feel an obligation to have children than when the reverse is the case.

D. Individual and Psychological Factors Related to Ideal Family Size

In addition to the reasons for having children which have already been reported, a number of other reasons were included, having chiefly to do with advantages accruing to the individual as a result of having children. Certain of these were considered very important by the group, especially those relating to effects on the family and the companionship and enjoyment of children. A few of these are given here followed by the percentages of officers and wives indicating that they regarded them as of "much" or "great" importance. *"Children tend to make the family more stable,"* officers—63 per cent, wives—64 per cent; *"For the companionship of young children,"* officers— 67 per cent, wives—68 per cent; *"Watching children grow up is a lot of fun,"* officers—53 per cent, wives—70 per cent; *"To provide a feeling of contentment in old age,"* officers —39 per cent, wives—52 per cent. A complete summary of the responses to Section *G* is given in Appendix *B*.

A significant relationship was found between the reports of ideal family size and the importance attached by the officers to having children because of enjoying them. In general very little relation was found in this group between reported "ideal family size" and scores on general measures of interests, values, and personality traits. The largest differences were found for the groups reporting "two" and "four" children as their "ideal family size" on the scales of the Allport-Vernon *Study of Values*. The scores of the group giving "four" as their "ideal family size" were significantly greater than those for the individuals reporting "two" with respect to the Religious Scale and significantly smaller with respect to the Political Scale for both officers and wives. Both officers and wives reporting the larger number as "ideal" were lower on the Economic and Aesthetic Scales than the corresponding groups reporting "two" as the ideal but the differences were so small that they could easily have been due to sampling fluctuations. Similarly there were no large differences with respect to these two groups on the Optimism-Pessimism Scale used or the scales of the Bernreuter *Personality Inventory* or the scales of the Strong *Vocational Interest Blank*. The attitudes, interests, and beliefs which influence the individual's concept of ideal family size seem to be quite specific and not par-

ticularly related to other general attitudes and personality characteristics.

In summary it appears that the individual's concept of ideal family size is largely determined by what the individual has heard or been taught. In some cases it is probably a "stereotype" with no very large amount of dynamic influence in practical situations.

IV. DETERMINERS OF FAMILY SIZE

A. Deviations from "Ideal" Family Size

In the preceding chapter the development of the individual's concept of the ideal family size has been discussed. The question now arises, *"Does the individual have the number of children he considers 'ideal' and if not what are the factors chiefly responsible for these deviations from this 'ideal' number?"* In answer to the first part of this question we find that whereas a majority of both officers and wives reported that they considered three or more children (average equal 3.24) as the "ideal," most of them expect to have two children or less (average equals 1.95) themselves. Both officers and wives attribute most of this discrepancy to the general conditions associated with life in the Army Air Corps. The average of the values reported as the ideal number of children for an Army Air Corps Officer is 2.53 for the officers and 2.44 for the wives. It is further apparent that they attribute much of this discrepancy to factors associated with all branches of Army Service, for the average of the values reported as the ideal number of children for Army Officers is 2.72 for the officers and 2.55 for the wives.

It should not be surprising that such a large part of this discrepancy is attributed to factors common to the group since so many factors such as salary, living conditions, social activities, and occupational hazards are very closely similar for all of them. The remaining part of the discrepancy is presumably to be accounted for by individual factors such as health, physiological conditions, marital happiness, and dependents other than wife and children.

Discrepancies between actual family size and ideal family size are not always in the same direction as is the case for this group. There are a number of individual cases in which the number of children is greater than the number considered ideal by the persons concerned. In some cases this may be attributed to the manner in which these data were reported but more frequently it appears to be due to a lack of control of family size. It is clear from the data presented in the preceding chapter that in certain other social groups, particularly at the lower income levels, the actual size exceeds the ideal size largely because of lack of adequate control of family size.

Before proceeding to a discussion of the factors which seem to be of greatest importance in determining actual family size, the problem of estimating size of completed family will be considered. The best

estimate of size of completed family is derived by utilizing data only from parents who are apparently beyond the child-bearing ages. Because of the severe restriction this would place upon the number of families available in this study and also because of the greater interest in the younger groups who are actually in process of determining family size, this group has been segregated only for certain checking purposes in this study. Another procedure which has been frequently used is to segregate groups of individuals of the same age who have been married the same length of time. The analysis of the data provided in this study yielded no significant differences between the number of children born to officers or wives who married at early ages and those who married later. The average number of children born in the first eight years of marriage to officers married between the ages of 20 and 24 was 1.67. In the same period officers married at ages 30 to 34 had 1.62 children. Wives show a slightly greater but insignificant difference. The average number of children born in the first eight years of marriage to wives married between the ages of 17 and 19 was 1.75. The corresponding figure for wives married between the ages of 25 and 29 is 1.54. Although age at marriage is thus seen to have little influence in this group the length of the marriage is a very important factor. For certain of the investigations made the number of children born during the first eight years of marriage was used. This seemed to be a very satisfactory index for this group since the product-moment correlation coefficient between this number and the number of children ever born to those 55 wives in the group who were over 40 years of age was found to be .93. It appears likely that this coefficient would not be so high for more heterogeneous groups. A similar correlation coefficient computed for the 148 sisters of the officers and wives was found to be .64. A part of this discrepancy is probably due to actual differences in the groups, but it also seems likely that it may be attributed partially to the group's lack of accurate knowledge of the marriage dates and ages of the children of their sisters.

However the imposition of even such a restriction as eight years of marriage did not appear desirable for the purposes of this study and therefore some other means of predicting size of completed family were sought. The estimate adopted for the main portion of this analysis was a report of the number of children already born plus the number of additional children expected. Although such an index

introduces a subjective element and a factor of uncertainty into the analyses, it seemed definitely the most satisfactory for this study.

The fact that the average of the figures reported as number of children planned (1.95) is fairly similar to the number of children born to those officers and wives whose families can be considered completed lends some support to the selection of these values. In general this evidence indicates that the values are probably a little high since unforeseen circumstances will prevent some families from having as many children as they plan. It would appear that in a few cases more children will be born than are planned but the replies to the questionnaire indicate that the younger members of this group are able to control the size of their families quite effectively.

Of the 238 couples in this study for which complete data for both husband and wife were available and not including cases involving remarriage or sterility, 202 husbands and wives agreed and 36 disagreed in their estimates of completed family size. In 20 cases the officer estimated that he would have one more child than his wife expected and in one case an officer estimated he would have two more children than his wife expected. In 13 cases the wife estimated that she would have one more child than anticipated by the husband and in two cases the wife expected to have two more children than estimated by the officer.

In the analyses which follow comparisons will be made of those planning none or one child who will be called the non-prolific group, with those planning three or more children who will be referred to as the prolific group.

B. Physical Factors Influencing Family Size

In this section certain fundamental factors such as sterility, birth control, and postponement of marriage will be discussed. In this sample there were 105 officers and 82 wives who have had no children and are not expecting one at present. Of these only 27 officers and 13 wives report that they do not want any children. In many cases these couples have not been married long and 47 of the officers and 38 of the wives report that they have made no effort to have children. However 52 of the officers and 37 of the wives state that their efforts to have children have been unsuccessful. Thus about 15 per cent of the officers and 13 per cent of the wives who have tried to have children have been unsuccessful. It is likely that a number of these

will have children later and therefore the proportion of involuntarily sterile marriages is probably somewhat below these values.

In 39 cases the officers reported that a physician had been consulted, but reports indicate that a definite diagnosis was made in only about half of the cases. In those cases where a specific defect was reported, about 80 per cent were attributed to the wives. In this group there did not appear to be any greater percentage of sterility for officers married at later ages than for those married much earlier in life. On the other hand 6 of the 17 wives married at age 28 or later reported that they had been unable to have children. This proportion is significantly different from the figures showing 19 of the 133 wives who were married at age 27 or earlier reporting sterility after at least eight years of marriage. It was reported in the previous section that there is practically no difference in the number of children born to officers in this group marrying early as compared with those marrying at a later age. A difference is found in the case of the wives, though the number of cases is not sufficiently large to provide a satisfactory degree of confidence in its existence in the larger group from which this sample was obtained.

From the evidence obtained in this study it would therefore appear that if the men in a professional group such as this marry relatively young wives, their own age at marriage, at least up to ages of around 35, will not affect the ultimate size of their family. This may be explained at least in part by the increasing tendency to control the size of family in such groups as this. In this group about 45 per cent report making definite plans concerning number of children shortly after marriage. Although the proportion making definite plans is larger in the younger groups, even in this latter group of married officers under 30 only 53 per cent report having made definite plans shortly after marriage. It appears that size of family is not yet generally regarded as something which may be planned in advance. That there is reason for this attitude is shown below.

Table 8 shows the number of accidental pregnancies (those in which conception occurred despite some attempted method of control), the number of definitely planned pregnancies, and the number which were not definitely planned but in which no preventive method was used. The pregnancies include only those which resulted in a normal birth and are classified according to the birth order of the child.

Although a significant difference does not exist between the pro-

TABLE 8

THE CLASSIFICATION OF PREGNANCIES ARRANGED BY BIRTH ORDER ACCORDING TO THE EXTENT TO WHICH THEY WERE DEFINITELY PLANNED AS REPORTED BY OFFICERS AND WIVES

	Pregnancy accidental		Pregnancy not definitely planned		Pregnancy definitely planned		Item omitted	
	O	W	O	W	O	W	O	W
1st Child	55	59	114	78	111	100	4	1
2nd Child	39	37	51	35	54	52	3	2
3rd Child	16	17	18	9	12	14	0	1
4th Child	9	7	6	5	2	1	1	0
5th Child	1	0	3	2	2	1	0	0
6th Child	2	1	1	0	0	0	0	0
Totals	122	121	193	129	181	168	8	4
Per Cent (of those responding)	25	29	39	31	36	40	—	—

portions of accidental pregnancies in the first and second born children, the proportion of accidents tends to become significantly larger for the later born children. A very significant finding which provides a partial explanation of this trend is that the officers over 40 years of age whose families are more frequently completed and who are responsible for many of the larger families report that only 20 per cent of their second born were planned, whereas the group of officers under 40 reported that 49 per cent of their second born were definitely planned. About 92 per cent of the group, excluding those reported as sterile, indicate that they have made use of some method of birth control. That these methods are not entirely effective for these individuals is indicated by the fact that even in the younger group 24 per cent of the pregnancies have been accidental, that is, they have occurred despite some attempted method of control.

In addition to collecting information on the negative side of birth planning, data were collected on the positive side which appears to have been even more completely neglected by previous studies. The question was asked, *"If the pregnancy was definitely planned, how many months after you stopped taking precautions did conception occur?* The median value of the responses given by the officers was 2.0 months. The corresponding value reported by the wives was 2.4 months. These data lend no support to the opinion voiced by some that the declining birth rate in this country is evidence of a general loss of virility due to the decadence of our national stock.

In concluding the discussion of the effect of these physical factors on family size, two general implications appear to deserve special importance in influencing family size and that both more extensive and intensive studies are needed of what might be called "temporary" and "permanent" sterility. More accurate information concerning the effect of wife's marriage age is especially needed. Second, a more intensive study of the changes in psychological attitudes accompanying the current trend towards a smaller proportion of "accidental" as compared with "definitely planned" pregnancies should be made. It would be particularly valuable to know whether increased control would lead to an increased amount of definite planning of family size prior to or shortly after marriage. For example, it might be that the lack of complete effectiveness of earlier methods of birth control tended to produce an attitude that the parents were likely to have a larger number of children than they wished in spite of their best efforts to the contrary and therefore that they should try to have as few as possible. The evidence obtained from Question 21 in Section *E* does not seem to point to any large effect on planned size as a result of an increased effectiveness of birth control methods.

C. The Influence of the Wife's Health and Related Factors on Family Size

Several types of information were obtained concerning the influence on family size of the wife's health and various circumstances affecting the wife's health. In Section *E* of the Family Report Form 29 per cent of the officers and 26 per cent of the wives report that they would plan to have a larger family if "Painless and safe childbirth were assured by advances in medical science." In response to another question in this section 69 per cent of the officers and 46 per cent of the wives report that they would plan to have a smaller family if "The wife could have children only by caesarian operation." It is apparent that the husband's concern for his wife's health is a definite factor in determining family size. Further evidence on this point is provided in Section *F*. More than a quarter of the officers state that consideration for the wife's health has been one of the factors preventing them from planning additional children.

Although these data suggest that the husband's concern for the wife's health is greater than her own, the wife's own fear of childbirth cannot be regarded as a negligible factor. More than 10 per cent of the wives who have not had children and are not now expect-

ing any report that they have been afraid of childbirth. In the group who have had children, 8 per cent report "much" fear of childbirth before the birth of the first child, 22 per cent report "some" fear, 64 per cent report "little," and 7 per cent report "none." The proportions in the various categories are very similar for later births. There appears to be very little difference between the groups planning small families and those planning large families with respect to their reports of the wife's fear of childbirth before the birth of the first child. The wives reported that of the recoveries from 402 births, 55 or about 14 per cent could not be considered normal recoveries. These included specific injuries sustained during childbirth, necessary operations following childbirth, and various complicating illnesses.

It should not be supposed, however, that this group regards the having of children as entirely negative in its effect on the wife's health. Forty-three per cent of the officers and 44 per cent of the wives report that the improvement of the physical and mental health of the mother is of "some," "much," or "great" importance as a reason for having children.

It is clear that the husband's consideration for the wife's health and the wife's fear of childbirth both have a definite effect on size of family. This effect is probably not a very large source of deviation from "ideal" size but it cannot be ignored in any analysis of the factors involved.

D. The Influence on Size of Family of Various Factors Related to Marital Status

It was pointed out in an earlier chapter that almost all of the officers in the Army Air Corps marry. The percentage seems to be distinctly higher than that for the general population. There were 38 single officers included in the sample obtained for this study. All but seven of these expect to marry sometime. The principal reasons given for not being married are that they have not found the right person and that they have had unusual financial obligations. These officers hope to marry "good health," "good disposition," and "brains." They report that "beauty," "money," and "social prestige" are of much less importance to them and they rank "good homemaker" in an intermediate position probably because of its generality.

The officers who are married indicate that the number of children

they expect to have is determined in part by the characteristics of their wives. Twenty-one per cent of the officers report that they would have a larger family if the person they married had exceptional ability to manage children well and to guide their personality development.

One of the reasons for having children rated as of most importance by both officers and wives is *"Being married to a person who wants children."* Eighty-three per cent of the officers and 90 per cent of the wives consider this reason to be of "some," "much," or "great" importance to them. Similarly 51 per cent of the officers and 47 per cent of the wives report that they would increase the number of children they expect to have if the person they married wanted a large family. Conversely, 68 per cent of the officers and 65 per cent of the wives would decrease the number of children they expect to have if the person they married did not want children.

Another factor which seems to be important is the happiness of the marriage. Sixty-four per cent of the officers and 60 per cent of the wives indicate that the number of children they expect to have would be reduced if married life was less happy than anticipated. There is practically no difference in the reported marital happiness of those planning large families as contrasted with those planning small families. However, marital unhappiness is reported by a small number of officers and wives as the reason for having no children or planning no more children. Furthermore, more than 93 per cent of the officers and 96 per cent of the wives report their marriages as "happy," "very happy," or "extremely happy." Officers and wives in the marriages not classified in one of these "happy" categories are expecting to have less children than the others but the number of cases is too small for the result to be statistically significant.

This group appears to be much more happily married than people in general. Forty per cent of the officers and 53 per cent of the wives report their marriages as extremely happy. There are probably many factors contributing to this such as the satisfaction of this group with their work, the regular employment and salary, and the general good health of at least the husbands in the group. A factor associated with marital happiness which has not received the attention in recent studies of this topic which it would seem to merit is the health of the wife. Of 48 wives who reported that they were never sick and in the best of health, not a single husband reported his marital happiness as below the classification "happy" and almost

two-thirds of these husbands said that their marriages were "extremely happy." Of 107 wives who reported that they were only occasionally ill, in good health, no husbands reported the marriages as "unhappy" and only six reported their marriages in the middle category of "neither happy nor unhappy." The largest group of these husbands reported their marriages as "very happy." Of 68 wives who described their health as "about average health" or "frequently ill, below average health," four husbands reported their marriages as "unhappy." It would seem likely that in a group in which the couples were less uniformly happy the use of more accurate measures of health and marital happiness would reveal an even larger degree of relationship.

It may be concluded from the discussion of this section that one of the most important factors in causing deviations in family size from the number the individual considers "ideal" is the characteristics, attitudes, and general compatibility of the person this individual marries.

E. General Social Factors Influencing Size of Family

It was mentioned in the previous chapter that social background is important in the formation of the individual's concept of "ideal family size." In this section the effect of various social customs, social pressure, and social conditions will be discussed. The bearing and rearing of children has long been regarded as a fundamental obligation of women who marry. Although it appears likely from our data that most women have children because they wish them, there appear to be a minority who have them because they feel obligated to or because they wish to conform to the current customs. Considerable social pressure appears to be exerted in some cases on individuals who would prefer to avoid the responsibilities of parenthood.

An important finding which appears to be associated with this factor is the relation between the relative socio-economic background and status of the husband and wife. By combining the individual's ratings of cultural and economic status of the parents' home with reports of the extent of the individual's own education and that of his parents, an index of socio-economic background was obtained. In most of the general unselected groups which have been studied in previous investigations those persons coming from the higher socio-economic levels have had smaller families on the average than those from the lower levels. However in this group there was a slight

tendency for the average estimated size of completed family to be larger for the officers with high socio-economic indexes than for those with low indexes. On the other hand the wives with high socio-economic indexes had an average estimated size of completed family lower than those with low socio-economic indexes. These findings will be discussed more fully in the next section.

The finding which is important to the discussion of this section is that wives married to officers whose social backgrounds as evidenced by this index are higher than those of the wives plan to have more children than wives married to officers whose social backgrounds are inferior to their own. This relationship while not large is sufficiently great to pass the usual tests for significance.

Evidence of a similar nature was obtained by studying the relation of outside income to estimated size of completed family. Wives with substantial outside incomes married to officers without appreciable outside income expected to have less children than the other wives. On the other hand wives with or without outside income married to officers with substantial outside incomes expected to have significantly more children than did other wives.

These findings support the hypothesis that women have children partly because of a feeling of obligation to their husbands, and when that feeling of obligation is diminished by outside circumstances the number of children planned by them is reduced.

Social activities in the Air Corps are relatively extensive. This may have some effect on size of family, but judging from the reports of the officers and wives this effect is small. Only 9 per cent of the officers and 8 per cent of the wives indicate that cutting the social activities of the Air Corps in half would cause them to increase the planned size of their families.

Indirect evidence concerning the factor of interference with social activities is provided by the reports concerning how much the wife worried about losing her figure after the birth of the first child. Twenty-eight per cent of the wives report "much" or "some" worry, while 40 per cent of the husbands report that their wives worried "much" or "some."

While pregnancy and the normal physical incapacity associated with childbirth appear to be of some importance, the care of the child after birth appears to be a more significant factor in influencing family size. The reports of the officers and their wives concerning

the way in which the care of their first child influenced their activities are given in Tables 9 and 10.

TABLE 9

RESPONSES OF OFFICERS AND WIVES TO THE QUESTION CONCERNING HOW THE CARE OF THEIR FIRST CHILD INFLUENCED THE HUSBAND'S ACTIVITIES

	Officer's own report	wife's report regarding officer
Made no difference	41	96
Stayed at home more	99	59
Had to curtail social activities	21	11
Had to rebudget expenses	14	22
Stayed at home more and had to curtail social activities	14	6
Stayed at home more and had to rebudget expenses	23	9
Curtailed social activities and rebudgeted expenses	11	13
Stayed at home more, curtailed social activities and rebudgeted expenses	39	13

TABLE 10

RESPONSES OF OFFICERS AND WIVES TO THE QUESTION CONCERNING HOW THE CARE OF THEIR FIRST CHILD INFLUENCED THE WIFE'S ACTIVITIES

	Wife's own report	Officer's report regarding wife
Made very little difference	110	104
Hampered social activities	87	132
Altered circle of friends	6	11
Stayed home more	4	5
Hampered social activities and altered circle of friends	12	7

Forty-four per cent of the wives report that the care of the first child made no difference in the husband's activities, and 50 per cent of the wives report that it made very little difference in their own activities. Only 16 per cent of the officers report that it made no difference in their own activities, and 40 per cent of them indicate that it made very little difference in the activities of their wives. The most frequently reported effects were that they stayed at home more and that their social activities were hampered.

One further bit of evidence on this point is obtained from the responses to one of the questions concerning the reason for a delay in having the first child. Almost 60 per cent reported that one of the

reasons for this interval was the desire for a period of freedom from the responsibility of children.

The second most frequent reason for limitation of family size in this group is the moving and living conditions associated with life in the Air Corps.

In summarizing this section it may be said that social background and customs and social conditions in the Air Corps undoubtedly cause deviations from "ideal family size" in this group. As was previously noted the social group also is important in determining this "ideal size." Although it is difficult to separate such indirect influence from the more direct effect related to specific social conditions, there can be no doubt that this latter factor does produce significant deviations from the size of family regarded as "ideal" by the individual.

F. The Influence of Economic Factors on Family Size

The factor which dominates a large part of the discussion of this group concerning family size and especially planned deviations from the size considered "ideal" is the financial one. In this group it is the more remote financial obligations which seem to cause the greatest concern. Education and insurance seem to be the factors which this group believe are the most important obstacles to having the somewhat larger families which most of them indicate they would prefer. In Section *E* 55 per cent of the officers and 49 per cent of the wives report that they would have increased the number of children planned if all the educational expenses of their children were provided for by a fund set aside by a "rich uncle." Similarly 53 per cent of the officers and 44 per cent of the wives report a planned increase would have resulted if a monthly income of $100 for the wife and $50 per month for each child to the age of 21 were provided by the government in case of the death of an officer.

This group reports that increased salary as well as specific benefits would result in increases in planned size of family. Forty per cent of both the officers and the wives report that if army pay schedules provided a 10 per cent increase in base pay at the birth of each child they would plan a larger family. About the same proportion, 40 per cent of the officers and 36 per cent of the wives, indicate that the increase of all pay schedules in the Air Corps by $50 a month would have a similar effect. That immediate expenses do not appear quite so important is shown by the fact that a somewhat smaller proportion, 29 per cent of both officers and wives, indicate that they

would have planned a larger family had $1000 been provided at the birth of each child by a fund set up by a "rich uncle." That the time of having these children did not seem very important to the group is indicated by the fact that only 8 per cent of the officers and 14 per cent of the wives indicated that an increase in planned family size would result from moving the first substantial increase in pay up one year earlier.

Evidence of the importance of the financial factor is also provided by the responses to a question involving a very drastic reduction in their effective incomes. Sixty per cent of the officers and 48 per cent of the wives report that if the purchasing power of the dollar were reduced to 30 cents they would decrease the planned size of their family.

It is interesting to note that in Section *F*, when questioned concerning the reasons for delay in having children, about one in three of both officers and wives reported that lack of money for the immediate expenses of having a child was a factor, and about one in four reported that uncertainty of ability to meet the necessary later expenses was a factor. However, in this same section about 60 per cent of each group report that their desire for a period of freedom from the responsibility of children was a definite factor. From this it appears reasonable to conclude that although financial factors are important in causing young couples to delay in having children, it is unlikely that this is the most important factor. This appears to be a significant finding for those interested in population problems.

Although we may conclude that the financial factor is probably not the primary factor in causing delay in having children, other responses in this section and in Section *H* indicate very clearly that this group regard it as by far the most important reason for not planning more children than they have. As mentioned in previous sections, other factors do operate in quite a number of individual cases but the proportions reported for such reasons are all small as compared with the various financial reasons which are given by more than half of the officers and wives.

In most of the reports given in the preceding paragraphs it will be noted that the officers tend to attach more importance to financial considerations than do their wives. This is probably to be expected since the financial responsibilities usually are the primary concern of the husband rather than the wife.

In Chapter II certain data were given concerning the rank and

TABLE 11

MEDIAN YEARLY FAMILY INCOME OF OFFICERS, WIVES, AND THEIR SIBLINGS BY AGE GROUP WITH PER CENT IN EACH GROUP HAVING A TOTAL FAMILY INCOME OF $10,000 OR MORE

	30-34			35-39			40 and over		
	N	Median	Per cent with incomes of $10,000 or more	N	Median	Per cent with incomes of $10,000 or more	N	Median	Per cent with incomes of $10,000 or more
Officers	153	4500	1.3	99	5500	2.0	105	7800	10.4
Officers' brothers	78	2200	0.0	98	2600	3.1	152	2900	7.2
Officers' married sisters	53	2600	0.0	56	3000	7.1	106	3200	6.6
Wives	91	4900	3.3	42	5800	4.8	51	7800	11.8
Wives' brothers	42	2900	2.4	40	5100	0.0	60	3200	5.0
Wives' married sisters	52	2600	1.9	44	3500	2.3	62	2800	1.6

salary of this group. Some of the individuals in the group had outside incomes which supplemented these salaries. A comparison of the total family incomes of the officers and wives with those of their brothers and sisters is given in Table 11. The values reported by the Air Corps group are very substantially greater than those which they report for their brothers and sisters. Although it should be remembered that the officers ordinarily have certain special obligations such as additional insurance premiums, their total annual family incomes appear to be relatively high.

Because of the findings reported in previous studies showing that individuals in the higher occupational and income groups tend to have less children than those in the lower levels, an analysis was made *by occupational status of the brothers and sisters of the officers* in this group. It was thought that an analysis of differences in birth rates by occupational levels within a group as homogeneous as this with respect to family and social background would be valuable. The results are shown in Table 12. When classified by occupational

TABLE 12

THE NUMBER OF CHILDREN BORN DURING THE FIRST EIGHT YEARS OF MARRIAGE TO THE FAMILIES OF THE BROTHERS AND SISTERS OF THE OFFICERS FILLING IN THE FAMILY REPORT FORM

	Children born in first eight years of marriage							
	0	1	2	3	4	5	Total	Mean
Professional	26	33	30	10	1	0	100	1.27
Proprietors, managers and officials	19	38	29	12	2	0	100	1.40
Farmers	4	13	9	2	1	0	29	1.41
Clerks and kindred workers	23	30	19	3	1	1	77	1.12
Skilled workers and foremen	9	9	7	4	0	0	29	1.21
Semi-skilled workers	3	5	6	0	0	0	14	1.21
Unskilled laborers	1	2	2	0	0	0	5	1.20

status of the husband no large differences are found. The average number of children born during the first eight years of marriage to the families of the brothers and sisters of the officers in which the occupation of the husband is reported as clerical is 1.12. This group has the lowest values. The similar groups classified as having managerial positions and as farmers provide the highest values. It is apparent that the trend usually found between occupational status and birth rate is conspicuously absent in this group. Similarly no

trend was revealed when the same type of comparison was made with respect to income.

A similar analysis of the figures reported for the brothers and sisters of the officers' wives failed to provide any definite indication of a trend. However, it was found that in this group the small number of farmers and the few individuals included in the group below the level of skilled workers had more children than did the families in which the husbands are classified in other occupational groups.

All groups combined for the siblings of the officers had an average value of 1.23 children during the first eight years of marriage; the average value for the wives' siblings is 1.41 children. These compare with the average values for these officers and their wives of 1.31 and 1.36 respectively.

Certainly these data indicate that the financial factor may hardly be regarded as a simple one. The incomes in this group are relatively high and yet they indicate that the greatest obstacle to increased family size is lack of money. The answer must be sought in terms of what this group consider their financial needs to be. Certainly most of them could maintain a standard of living far above that which they experienced in their parents' homes and still have plenty of money to raise several children. It is apparent that they feel that a standard of living far above that is of more importance to them than a large family. To analyze what this standard of living comprises, the proportions of family income spent for various items as reported by the officers are given in Table 13.

From an inspection of this table it is seen that the three large items, each of which accounts for approximately 15 per cent of the

TABLE 13

THE MEAN VALUES OF THE PERCENTAGES OF THE TOTAL ANNUAL FAMILY INCOME SPENT FOR VARIOUS ITEMS AS REPORTED BY THE OFFICERS

Rent	15.9	Recreation	4.0
Food	15.7	Hobbies	1.4
Clothing and its upkeep	9.0	Insurance and other savings or investments	14.9
Household furnishings, etc.	4.4	Payment on debts	3.4
Social activities	5.2	Travel (and car, excluding official)	8.0
Education (books, magazines, etc.	2.3	Contributions for support of dependents other than wife and children	3.2
Charity and church	1.1	Miscellaneous	5.0
Medical care	1.2		
Maid service	5.4		

total expenditures are rent, food, insurance, and other savings and investments. In this group as in other professional groups the proportion of the total income spent on "necessities" is very much smaller than in the lower income groups.

To study the effect of the financial factor on limitation of family size, the individuals were asked to indicate the changes in the amounts being spent on the various family budget items which would be made if one child were added to the present family. These values are reported in Table 14.

TABLE 14

ESTIMATES OF 166 OFFICERS AND 117 WIVES OF EFFECT ON FAMILY EXPENDITURES FOR VARIOUS ITEMS OF THE ADDITION OF ONE CHILD TO THE FAMILY

Item	Mean of officers' estimated changes	Mean of wives' estimated changes
Food	$122	$121
Clothing and its upkeep	76	79
Household furnishings, etc.	14	18
Social activities	—34	—25
Educational	33	22
Charity and church	0	0
Medical care	50	55
Maid service	71	56
Recreation	—23	—15
Hobbies	—8	—7
Insurance and other savings	30	40
Payment on debts	—13	—10
Travel (and car)	—29	—26
Contributions for support of dependents other than wife and children	—11	—7
Miscellaneous	—8	7

Since living quarters are ordinarily furnished to officers and their families, this item was not included. The items for which both officers and wives reported the largest increases were food, clothing, medical care, and maid service. The decreases indicated were not as large as the increases. The largest decreases were for social activities, travel, and recreation. These increases do not, of course, reflect the true cost of the child because within some of the categories part of the new child's portion would be taken from the amounts previously spent for these items by other members of the family. Furthermore they seem to indicate the more immediate changes in expense to a greater extent than those which are more remote such as education.

These considerations indicate that the total cost to these families of raising an additional child would be, on the average, at least

$17,500. Estimates of the way in which this amount might be spent are shown in Table 15. These estimates are probably very conservative and it might be more proper to regard them as "bare minimums" than average values.

TABLE 15
ESTIMATED MINIMUM COSTS TO AIR CORPS OFFICERS' FAMILIES OF RAISING AN ADDITIONAL CHILD

Food	per annum	$200.00
Clothing	per annum	100.00
Household furnishings	per annum	25.00
Education, excluding college	per annum	50.00
Medical care	per annum	10.00
Maid service	per annum	100.00
Insurance	per annum	200.00
Miscellaneous	per annum	65.00
	TOTAL per annum	750.00
	TOTAL for 18-year period	13,500.00
	College Education	4,000.00
	GRAND TOTAL	17,500.00

This group like most other professional groups can be considered *as fairly well-paid. Certainly their salaries are substantially greater* than those of their parents and brothers and sisters. Yet they consistently report that the most important reason for family limitation is lack of money either immediate or anticipated. It appears that they consider a high standard of living essential. It may well be asked whether the financial reason as reported really exerts any influence or is just a plausible statement which has been adopted by many of the group. Certain direct evidence has been obtained on this point.

In Table 16 an analysis of the births in this group is presented in terms of the year of service in which they occurred. The salary schedules of the United States Army call for an increase in pay every three years after the acceptance of a commission in the regular army. In Section *F* of the Family Report Form the officers were asked whether each pregnancy was intentional or not.

Inspection of this table reveals that most conceptions occur shortly after the officer has received an increase in pay, and fewest in the year preceding the next increase. This finding is very pronounced except in the case of the first, second, and third years of service. A

TABLE 16

AN ANALYSIS OF ACCIDENTAL AND INTENTIONAL PREGNANCIES ACCORDING TO THE YEAR OF COMMISSIONED SERVICE OF THE OFFICER IN WHICH THE BIRTH OF THE CHILD OCCURRED

	Year of commissioned service																	
	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	12th	13, 16, etc.	14, 17, etc.	15, 18, etc.	1, 4, etc.	2, 5, etc.	3, 6, etc.
Definitely planned	5	6	13	15	20	8	20	23	9	10	6	6	6	6	0	56	61	36
Not definitely planned	4	17	14	18	15	11	10	7	9	9	12	4	7	8	3	48	59	41
Accidental pregnancy	0	13	15	11	11	7	11	3	6	4	4	1	4	4	5	50	35	34
Total	9	36	42	44	46	26	41	33	24	25	22	11	17	18	8	154	155	111

different relationship during this period is to be expected since officers are in almost all cases single at the time they receive their commissions. It is interesting to note that if the first three years are disregarded a similar tendency is observed in both the group of pregnancies "not definitely planned" and the "accidents." It would seem therefore that "accidents" are not purely chance happenings but are more likely to occur when the financial circumstances of the family are relatively favorable.

Other evidence that the financial factor actually operates is obtained by comparing the size of the families of those officers with outside incomes with those of officers lacking such incomes. Officers with outside incomes plan significantly larger families than officers who do not have such an additional source of financial assistance. A possible explanation of the finding that wives who have an independent income but whose husbands have no outside income plan fewer children than the other wives of the group was suggested in the previous section. When both officers and wife have outside incomes the planned family size is also larger than in families having no such outside income.

The final evidence that the financial factor is an important one is provided by the data presented in the first chapter of this report. In that discussion it was reported that a significant increase in the number of fourth and later born children resulted from the offering of educational scholarships to such children of U. S. Army Air Corps Officers. The plan originally proposed was a combination of insurance to cover child care with aid for educational expenses. Such a combination might have been even more effective had it been possible to offer it. Unfortunately the international situation in 1939 prevented the insurance companies from making commitments to write such policies as far as a year in advance of the date that the offer had to be announced.

It must be clear from the foregoing discussion that the financial factor is one of definite importance and actually operates in the determination of family size in a group in which family limitation is generally practiced. Evidence to the contrary which has been frequently reported showing smaller birth rates in the higher income groups, appears to be derived from studies on populations in which part practiced birth control and part did not.

G. The Influence of Uncertainty of the Future on Family Size

The final factor which will be discussed is closely related to the

financial factor in many respects. This is the influence of uncertainty of the future. In addition to the general factors of uncertainty which confront most people, this group has the additional factor of the hazards of military flying. Evidence of the effect of this factor is provided by the responses to two items in Section *E*. If accidents in the Army Air Corps were no more numerous than those of civil life, 26 per cent of the officers and 30 per cent of the wives report that they would plan larger families. On the other hand 47 per cent of the officers and 44 per cent of the wives indicate that if flying hazards in the Army Air Corps were doubled by more complex equipment they would *reduce* the planned size of their family.

In Section *C* the officers report the values of their estates in the event of aviation accident. The median value is just a little below $25,000. This is mainly in the form of insurance benefits. The officers may take out a $10,000 policy of government War Risk Insurance at the time they are commissioned. For this insurance, rates are quite reasonable. On all additional insurance, however, the officer is required to pay an extra premium for aviation risk amounting to about $10 per year per thousand dollars worth of insurance.

These officers receive 50 per cent additional base pay when ordered on flying status providing they fly a certain minimum amount. Most Air Corps Officers qualify for this additional pay. Although this extra flying pay enables them to obtain additional insurance, no amount of insurance can adequately compensate for the loss of the head of the family.

Certain evidence on the influence of flying hazard on size of family is provided by comparing the birth rates in other branches of the United States Army with those in the Air Corps. Data for such a comparison is given in certain unpublished tables supplied by the War Department. The data from two surveys, one for 1928 and the other for 1933 are reported in Table 17, classified according to three pay period groups.

Although the samples from the Air Corps are too small to provide a precise comparison of the two groups, no very marked differences in children per officer are evident. This is of considerable interest since the social backgrounds, general living conditions, and occupations of the two groups are very similar. There are two conspicuous differences in the groups. First, the factor of flying hazard operates in the Air Corps group and has no very direct counterpart in the other branches; and second, the Air Corps group receives an amount

TABLE 17

NUMBER OF CHILDREN IN FAMILIES OF OFFICERS IN THE U. S. ARMY AIR CORPS AND IN ALL OTHER BRANCHES OF THE ARMY CLASSIFIED BY PAY PERIODS

	Year of survey	Fourth pay period Air Corps	Fourth pay period Other branches	*Third pay period* Air Corps	*Third pay period* Other branches	*Second pay period* Air Corps	*Second pay period* Other branches
Number in group	1933	18	503	108	1,044	45	350
	1928	12	358	88	942	34	275
Average age	1933	44-9	44-10	39-4	39-4	32-1	31-7
	1928	42-2	41-5	35-8	36-4	32-8	30-10
Number of children	1933	36	846	150	1,663	50	401
	1928	24	594	101	1,362	47	265
Children per officer	1933	2.00	1.68	1.39	1.59	1.11	1.15
	1928	1.68	1.66	1.15	1.45	1.38	.97

equal to 50 per cent of their base pay for flying duty. It is interesting to observe that there appears to be some tendency for the effects of these two factors to counteract each other in the Air Corps group.

It is possible that the fact that the Air Corps officers have almost as many children as the officers in other branches of the service is the result of a generally optimistic point of view concerning the future in general and their own personal welfare in particular. Evidence favoring this interpretation is found in their responses to the items on a scale of optimism-pessimism. The Air Corps officers appear to be definitely optimistic and there is a tendency for the most optimistic of them to plan slightly larger families than the others.

V. SUMMARY AND CONCLUSIONS

This study represents a somewhat different approach to the problem of analyzing the factors determining family size than that represented by previous studies. Instead of studying the birthrates in a few large social, economic, or geographical groups, over two thousand separate items of information were obtained concerning a relatively small and homogeneous professional group consisting of about 400 families. This study differs not only in the amount but also in the type of information obtained from each family. Most previous studies have been confined to the collection, tabulation, and analysis of objective data such as age, race, education, religion, occupation, income, size of community, and age of marriage. The present investigation, in addition to such factors, includes many subjective and psychological factors such as attitudes, interests, preferences, and values.

The procedure selected for obtaining the data in this study was a modified questionnaire procedure in which the general aims and methods of the study were explained to the men, either in groups or individually, and they were then given two sets of forms, one set to be filled in by themselves and one by their wives. The forms were prepared by a committee and tried out with two or three groups and carefully revised before being used in this study.

The group from whom the data reported in this study was gathered consisted of a sample of 427 officers in the United States Army Air Corps and 320 of their wives. The data was obtained by a flight surgeon on leave from the United States Army Medical Corps who visited most of the large Air Corps fields in the country. A very large proportion of those to whom the problem was presented filled out the forms even though this usually required more than three hours. Probably the factors of most importance in gaining the coöperation of the group were the intrinsic interest of the problem, the precautions to ensure anonymity, the offer of a report on the personality and interest blanks, the official approval which had been given the project, and the possibility of obtaining a vocational interest pattern on the Strong *Vocational Interest Blank* which might be of assistance in the selection of future officers.

The sample obtained included almost one-third of the regular officers in the Air Corps. The sample contained a greater proportion of the younger officers, those with lower ranks, and those who are

graduates of the United States Military Academy than would have been the case in a truly representative sample. However, it included officers from all of the various broad categories of rank, age, and family status. It does not appear that the method of sampling or the procedures used in collecting the data introduced a serious bias which might invalidate the general findings of the study.

It was found that although an unusually large proportion of the officers marry, this group was falling far short of having enough children to reproduce themselves.

One of the aims of this study was to analyze the factors involved in the formation of a concept of ideal family size. The problem is essentially one of determining how many children an individual would have if such limiting factors as health, restriction of activities, financial resources, characteristics of spouse, and the uncertainties of life were made ideal. With such physical and environmental factors removed the problem becomes a purely psychological one dependent on personality traits, interests, attitudes, values, and beliefs.

Approximately 95 per cent of both officers and wives reported that they regarded two, three, or four children as the ideal number for the average American family. The average of the values reported, 3.24 for both officers and wives, agrees very well with the findings which have been reported for practically all different types of groups. Two findings are especially important. First, there is very little difference between a highly selected professional group such as this and groups such as those at the other end of the socio-economic scale who are "on relief" in the average number of children regarded as *ideal*. Second, there is a marked tendency in this as in other groups for the youngest individuals to report as their ideal number of children a smaller value on the average than the remainder of the group. It would be valuable to know whether the individual's concept of ideal size gets larger as he grows older and more mature or whether we may expect the next generation to grow older with a markedly different idea of ideal size than has characterized the present generation.

A summary is given of the extent to which the concept of ideal family size is traceable to each of three types of factors: first, the conditions and training of the individual's immediate family group; second, the corresponding aspects of the wider social group of which the individual is a part; and third, those individual and personal characteristics which cannot be readily attributed to any particular environmental influences:

(*a*) The individual's concept of ideal family size is essentially an attitude, and like most other attitudes can be expected to have been considerably influenced by factors associated with the parent's home and family. It would be interesting to compare the ideal number of children reported by persons from the same family to determine the proportion of the individual variation in this concept due to factors related to the home environment. Although data were not available to enable an exact determination of the influence of this factor, the importance of the factor was definitely established. It was found that the chances are more than two to one that an officer who reports four or more as the ideal size of family comes from a family of four or more children, whereas among officers who report two or less as the ideal size the chances are greater than two to one that the officer comes from a family of three or less children. There is a slight tendency for the individual's statement of ideal family size to be positively associated with the individual's ratings of parent's marital happiness and negatively related to ratings of childhood happiness. These relationships are too small to be significant but might be investigated with a larger group.

Religion is a factor closely associated with the home which has been found related to family size by previous investigators. In this group the particular denomination to which the individuals belonged made little difference in the reported ideal size of family. On the other hand "extent of early religious training" was found to be significantly related to ideal family size for the wives and "church attendance" was significantly related to the officer's report of ideal size of family. Scores on the Religious Scale of the Allport-Vernon *Study of Values* were also found to be significantly related to the individual's reported ideal number of children. On the other hand only a small portion of the group considered that an important reason for having children was that it was a religious obligation. Although some of them considered it important to have children to carry on the family name and social traditions, very few of them attached much importance to carrying on the family enterprises or professional traditions.

(*b*) The second large group of factors which contribute to the concept of ideal family size are the social factors. More than half of the group attach much importance to the statement that *"Good stock ought to reproduce itself"* as a reason for having children. However, very few of them regard a general social obligation for *all* individuals to have children as of much importance to them.

The responses to other questions reveal that the predominant reason for having children in this group is that it is the natural and expected thing. That the wives generally feel some obligation to have children for their husbands is shown by the fact that when the financial dependence of the wife on the husband is decreased by virtue of her having greater financial resources than he and a substantial independent income, she plans to have less children than when her financial dependence is greater.

(c) A part of the individual's concept of ideal family size appears to be due to individual factors which cannot be readily attributed to any particular environmental influences. The fact that some individuals enjoy the company of children to a greater extent than others appears to be due at least in part to a difference in temperament. A definite relationship was found between the amount of importance attached by the officers to enjoying children as a reason for having them and ideal size of family.

The attitudes, interests, and beliefs which influence the individual's concept of ideal family size seem to be quite specific and not particularly related to other general attitudes and personality characteristics as measured by available tests. The individual's concept of ideal family size is largely determined by what the individual has heard or been taught.

It was found that only a rather small proportion of the individuals in this group have the number of children which they indicate as ideal. Whereas most of them indicate that they consider three or more children as the ideal number, the majority are planning two children or less themselves. Both officers and wives attribute this discrepancy to the general conditions associated with life in the Army Air Corps. In certain other groups it should be noted actual size of family exceeds ideal size largely because of lack of adequate control of family size.

For the analyses of this study two estimates of size of completed family were used. The first was number of children born in the first eight years of marriage which correlated .93 with number of children ever born to those wives in the group over 40 years of age. The second, which was used in most of the analyses including the summary analyses reported in Appendix B, was a report of the number of children already born plus the number of additional children expected. Those expecting to have none or one child in the completed family

are called the non-prolific group and those planning three or more children are called the prolific group.

The problem was to study the influence of various physical and environmental factors which cause the actual sizes of families to vary from those reported as ideal. The findings for several types of factors are summarized below:

(*a*) Physical factors have considerable direct importance in influencing family size. Although age at marriage appeared to have little effect on the size of officers' families, a significantly greater proportion of cases of sterile marriages is noted for wives who married at ages greater than 27 when compared with the remainder of the group. A more accurate investigation of this finding would be desirable. Another finding which appears deserving of more careful study is contained in the responses to the question, "*If pregnancy was definitely planned, how many months after you stopped taking precautions did conception occur?*" The median value of the responses given by the officers was 2.0 months and the corresponding value reported by the wives was 2.4 months. These data lend no support to the opinion voiced by some that the declining birth rate in this country is evidence of a general loss of virility due to the decadence of our national stock. A thorough investigation of this problem of "temporary" sterility would be valuable.

A definite trend toward a larger proportion of "definitely planned" pregnancies was observed. For officers over 40 years of age only 20 per cent of the second born children were reported as "definitely planned," while for officers under 40 years of age 49 per cent of the second born children were so reported.

Only 45 per cent of this group report making definite plans concerning number of children shortly after marriage. A more intensive study of the changes in psychological attitudes accompanying the current trend towards a smaller proportion of "accidental" as compared with "definitely planned" pregnancies should be made.

(*b*) The husband's consideration for the wife's health and the wife's fear of childbirth both play a definite but relatively minor part in determining size of family. It should also be noted that almost half of both the husbands and the wives report that the improvement of the physical and mental health of the mother is of some, much, or great importance to them as a reason for having children.

(*c*) Factors related to marital status, such as the characteristics, attitudes, and general compatibility of the person the individual mar-

ries, appear to be very important in causing deviations in family size from the number the individual considers ideal. Such a large proportion of the husbands and wives in this group reported their marriages as very happy that direct evidence concerning the effect of this factor was quite meager. However the statements of the officers and wives indicate that they consider it of great importance. An interesting related finding brought out by this study is the importance of the wife's health for marital happiness. This finding appears to have been neglected in recent studies on this topic. It is likely that in a group in which the couples were less uniformly happy, the use of more accurate measures of health and marital happiness would reveal an even larger degree of relationship than that found in this study.

(*d*) General social factors such as social background, custom, and social conditions undoubtedly cause deviations from ideal family size in this group. A significant tendency is observed for wives married to officers whose social backgrounds are higher than those of the wives to plan to have more children than wives married to officers whose social backgrounds are inferior to their own. Evidence of a similar nature was obtained by studying the relation of outside income to estimated size of completed family. These findings support the hypothesis that women have children partly because of a feeling of obligation to their husbands, and when that feeling of obligation is diminished by outside circumstances the number of children planned by them is reduced.

Moving and living conditions are mentioned more frequently than any factor except the financial one as a reason for limitation of family size by this group. Almost 60 per cent of the group who reported an interval of 18 months or longer from the date of marriage to the birth of the first child gave as one reason for this delay the desire for a period of freedom from the responsibility of children.

(*e*) The economic factor dominates a large part of the discussion of this group concerning planned deviations from the size of family considered ideal. Educational expenses and insurance are reported as the most important obstacles to having the somewhat larger families which most of them indicate they would prefer. A finding with rather general implications for population policy is that although the financial factor is important in causing young couples to delay in having children, it is probably not nearly as influential as the factor mentioned above concerning desire for a period of freedom from the responsibility of children.

In reporting on reasons for limitation of family size the husbands attach more importance to financial considerations than do their wives. The family incomes of the officers are conspicuously higher than those of their brothers and sisters but the number of children born to officers in the first eight years of marriage is slightly lower. Furthermore when classified by occupational or income level groups, these siblings do not show the marked differences in birth rate usually found in occupational groups.

An analysis of budget expenditures indicated that like other professional groups these individuals spend a much smaller proportion of their income on "necessities" than the lower income groups. It is evident from these reports and those concerning the cost of an additional child that the financial factor is a very complex one which must be related to standards of living in the group being studied.

Three pieces of evidence are reported which show that the financial factor really operates in this group and is not merely a plausible statement which has been adopted for convenience. First, it is shown that "definitely planned" births and to some extent others also are clearly related to the years in which an officer receives an increase in pay. This is at the end of every third year in the Army Air Corps. Second, officers with outside incomes plan significantly larger families than officers who do not have such an additional source of funds. Third, the final evidence that the financial factor actually operates is provided by the fact that a significant increase in the number of fourth and later born children resulted from the offering of educational scholarships to such children born to regular United States Army Air Corps Officers during the calendar year 1940.

(*f*) Uncertainty of the future and particularly the hazards of military flying appears to have a definite influence on the planned sizes of families in this group. The fact that the group have almost as many children as officers in other branches of Service appears to be due to a tendency for the effect of flying pay to counteract the effect of flying hazard in this group.

In conclusion it is repeated that the present study makes no pretense of untangling the complex matrix which includes the factors responsible for the determination of family size. It is hoped that it will prove of value in indicating useful procedures and pointing out which of the many types of information obtained in this study should be found most profitable in future studies.

APPENDICES

Appendix A: Announcement of Educational Scholarships

The announcement of educational scholarships sent in April, 1939, to all regular officers of the United States Army Air Corps having at least two living children is reproduced below.

Announcement of Educational Scholarships

Available for the Benefit of Children of Officers of the United States Regular Army Air Corps Under the Conditions Outlined Below. (From January 1 to December 31, 1940.)

Conditions: *THE PIONEER FUND, INC.*, 31 Nassau Street, New York City, offers fully paid-up annuities to provide educational scholarships for children of officers of the United States Regular Army Air Corps, as follows:

Any child born to an officer of the United States Regular Army Air Corps listed by the Official Army Register of January 1939, and to whom The Pioneer Fund, Inc. has mailed this Announcement, will be granted such an annuity when the following conditions are met:

1. The child who is to be the beneficiary of the grant must be born between January 1, and December 31, 1940, inclusive.

2. The officer in question must have three or more other living children at the time this particular child is born (or at some time after May 1, 1939).

3. Application for the grant must be made within a reasonable time after the birth of the child and not later than April 1, 1941.

Cost: These annuities will be provided at *no* cost to the officer. They will be issued in the form of a regular single premium annuity contract written in the child's name by a nationally known insurance company, on which the entire premium will be paid The Pioneer Fund, Inc. at the time of issue. The contract will be placed in the hands of the officer at that time and The Pioneer Fund, Inc. will have no rights or interest therein after payment by it of the premium. Thus it will constitute an outright gift of a fully paid-up annuity on which no premiums will ever have to be paid by the officer.

Provisions of the fully paid-up educational annuity: The Pioneer Fund, Inc. will place a sum in an annuity for each qualified child which will provide:

1. *A total of four thousand dollars toward the child's maintenance and educational expenses*: This sum will be paid in installments of five hundred dollars yearly for eight years,

beginning at the end of twelve years from the date of issue of the annuity; i.e., payment will start soon after the child's twelfth birthday (assuming that application for the grant is made soon after birth) and conclude soon after the nineteenth.

2. *In the event of the child's death* before the above payments begin, a lump sum payment of between two thousand seven hundred and three thousand six hundred dollars (depending on the child's age at death) will be made to the parents for the maintenance and education of the surviving children.

Applications: To apply for such a grant the officer should write as soon as possible after the birth of a qualified child, to Mr. John H. Slate, Jr., 31 Nassau Street, New York, N. Y., who is the legal representative of the Board of Directors of The Pioneer Fund, Inc. Should the officer die before having made such application, or if it becomes otherwise impossible for him to do so, his wife or other person responsible for the child may apply in his stead. Application blanks and further information will be furnished at any time upon request.

Every child meeting the conditions stated above will be automatically qualified for a scholarship grant; there will be no further conditions or requirements. Funds sufficient to provide scholarships for more than the full number of such children who could reasonably be expected in the calendar year 1940 are available for this purpose.

Appendix B: A Tabular Summary of the Scores and Responses of Officers and Wives for the Various Forms and Schedules Used

This summary gives the scores and responses of officers and their wives for the various forms and schedules used. Condensed distributions and means are provided. The code used in reporting the distributions is explained. Means for certain of the items have also been reported for the two groups who indicate that they regard two children and four children respectively as the "ideal number of children for the average American family." These two groups contain 79 and 123 of the officers and 67 and 116 of the wives in the two and four children groups respectively. Similarly means have been provided for some of the items for the groups which have been designated as non-prolific and prolific. Individuals are included in the non-prolific group who expect to have none or one child and those who expect to have three or more children are classified in the prolific group. These groups contain 125 and 116 of the officers and 90 and 83 of the wives in the non-prolific and prolific groups respectively.

The reports of codes proceeding by regular step intervals are abbreviated. In all of the codes X indicates the number of individuals who failed to report on the particular item. In those codes in which 0=*no* and 1=*yes*, the mean may be interpreted as the proportion of "*yes*" responses.

The directions printed on the cover page of the Family Report Form are quoted in full below:

> The War Department and Administrative Staff of the Army Air Corps are coöperating in a study of family life. The study is voluntary and anonymous; no name will be associated in any way with any blank you fill out. You will place your blank in an envelope and seal it. It will be sent directly to the Statistical Bureau of Columbia University for analysis. No blank will ever be seen by anyone at this post, or at any post, or by anyone in the War Department. The results will be analyzed and a statistical summary prepared. No statistical data will be published without the full approval of the Office of the Chief of the Air Corps and the War Department.

FAMILY REPORT FORM

Section A: Your Parents

	Item	Code	Officer or Wife	Distribution 0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
1c	Father's education	0=8 years or less	O	107	35	100	56	62	48		19	12.14				
		1=9 or 10 years	W	46	21	64	53	64	40		32	13.18				
		5=17 years or more														
1d3	Father's siblings	0=0	O	13	57	51	82	53	65	117	9	4.32	3.76	5.05	4.07	4.86
		1=1	W	19	33	40	47	44	39	85	13	4.22	4.21	4.56	4.26	4.12
		6=6 or more														
1e1	Father's age at marriage	0=20 and under	O	17	61	75	73	54	66	70	11	26.77	25.31	26.76	25.65	26.98
		1=21 and 22	W	13	52	46	51	46	44	50	18	26.77	27.58	26.62	28.15	26.65
		6=31 and over														
1e2	Parents divorced	0=no	O	371	29						17	.07				
		1=yes	W	264	20						26	.07				
1e3	Parents separated	0=no	O	322	13						64	.04				
		1=yes	W	214	11						76	.05				
2c	Mother's education	0=8 years or less	O	89	37	142	83	52	4		19	11.82				
		1=9 or 10 years	W	45	18	119	71	37	4		26	12.30				
		5=17 years or more														
2d3	Mother's siblings	0=0	O	23	36	66	57	61	48	122	14	4.24	4.12	4.99	4.31	4.35
		1=1	W	17	27	45	42	44	41	90	14	4.36	4.26	4.31	4.05	4.00
		6=6 or more														
2e	Mother's age at marriage	0=20 and under	O	178	95	51	42	21	13	13	14	21.94	21.18	21.83	21.05	22.37
		1=21 and 22	W	126	64	48	34	16	10	4	16	21.64	22.15	21.51	21.82	21.82
		6=31 and over														
3	Parents' marital happiness	0=extremely happy	O	75	100	155	43	50	8	6	10	1.76	1.91	1.63	1.80	1.73
		6=extremely unhappy	W	67	57	95	55	26	7	10	5	1.92	2.20	1.59	1.83	1.85

Section A (*continued*)

	Item	Code	Officer or Wife	0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
4	Happiness of own childhood	0=extremely happy 4=extremely unhappy	O	141	205	54	25	0			2	0.91	0.82	0.94	1.11	0.92
			W	146	126	21	19	5			3	0.77	0.68	0.83	0.86	0.77
6	Parents feeling: number of children	0=too many 2=too few	O	19	263	16					129	0.99	1.02	1.01	0.91	1.01
			W	27	168	14					111	0.94	1.00	0.93	0.92	0.92
7	Own feeling: number of children	0=too large 2=too small	O	84	295	44					4	0.91	0.93	0.85	0.74	0.97
			W	53	216	41					10	0.96	0.95	1.01	0.91	1.04
8d	*Economic* status of parents when individual of college age	0=very poor 3=comfortable 6=wealthy	O	3	61	127	132	66	15	2	21	2.62				
			W	0	28	76	85	70	31	4	26	3.04				
8e	Economic status of self at present	same as 8d	O	0	2	43	221	150	9	0	2	3.28			3.29	3.31
			W	1	0	26	150	130	7	1	5	3.37			3.38	3.44

Section B: Your Parents' Children

	Item	Code	Officer or Wife	0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
2	Number of brothers and sisters	0=0 1=1 6=6 and over	O	22	77	100	79	54	37	58	0	4.21	3.12	5.13	4.21	4.83
			W	29	77	76	40	52	22	45	1	3.86	3.20	4.50	3.64	4.29

Section C: Yourself

Item	Code	Officer or Wife	Distribution 0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
1b *Age last birthday*	*0=24 or under*	O	5	77	158	87	54	31	15	0	35.07			37.53	35.21
	1=25 to 29	W	47	97	87	42	28	15	2	2	31.41			33.31	31.49
	6=50 or over														
2a Your education—number of years of schooling	0=11 years and less	O	0	15	48	152	148	50	12	2	17.46	17.54	17.33	17.40	17.25
	1=12 and 13 years	W	15	109	116	69	6	0	1	4	13.97	13.95	13.94	13.89	13.82
	6=22 years and more														
2b Level of schooling—West Point training	0=no West Point	O	250	175						2	.41			.36	.33
	1=West Point														
2d Was college coeducational	0=no.	O	161	198	23					6					
	1=yes	W	67	110	10					11					
	2=both														
4 Choice of vocation (with wife's choice for husband)	0=other	O	142	52	186	27				20					
	1=army	W	85	57	116	16				46					
	2=army air corps														
	3=aviation														
5a Age when receiving first army commission	0=20 and under	O	9	127	125	87	47	18	9	4	24.09	24.40	24.25	24.55	24.58
	1=21 and 22														
	6=31 and over														
5b Total years of service	0=0 to 6	O	87	179	69	78	9	4		1	1.72				
	1=6 to 12														
	5=30 to 36														
6 Rank	0=General	O	1	7	10	61	111	180	55	2	4.43			4.08	4.44
	3=Major														
	6=Second Lieutenant														
7 Number of years of active flying	0=0 to 3	O	19	47	68	115	55	31	87	5					
	1=3 to 6														
	6=18 and more														
11a Church membership	0=no	O	81	338						8	.81				
	1=yes	W	37	273						10	.88				

SECTION C (*continued*)

Item	Code	Officer or Wife	Distribution 0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
11b Church membership (Catholic and non-Catholic)	0=non-Catholic 1=Catholic	O	385	34						8	.08				
		W	267	45						10	.14				
12a Early religious training	0=very extensive 2=moderate 4=practically none	O	43	136	224	17	7			0	1.55	1.61	1.49	1.56	1.56
		W	43	99	155	13	7			3	1.50	1.62	1.33	1.51	1.46
12b Regular church attendance now	0=no 1=yes	O	392	34						1	.08	.05	.11	.04	.14
		W	270	48						2	.15	.08	.18	.12	.25
14 Officers' outside annual income (and wife's estimate of this)	0=0 1=1 to 501 6=2501 and over	O	327	64	15	1	2	2	5	11	$105				
		W	225	38	7	2	1	1	5	43	$110				
15 Wife's outside annual income of her own (and officer's estimate of this)	0=0 1=1 to 501 6=2501 and over	O	319	30	11	5	2	0	6	12	$100				
		W	241	22	13	2	1	1	10	30	$150				
16a Officer's net worth *or* net indebtedness (and wife's estimate of this)	0=more than $1000 in debt 3=about even 6=net worth about $5000 or more	O	19	15	28	52	31	118	159	5	$800				
		W	13	13	26	54	15	59	96	44	$650				
16b Wife's net worth or net indebtedness (and officer's estimate of this)	0=more than $1000 in debt 3=about even 6=net worth about $5000 or more	O	0	1	0	308	28	30	47	15	$670				
		W	0	1	5	197	23	29	39	28	$750				
18b Worth of estate—aviation accident (by *thousands*)	0=under 10 1=10 to 20 6=60 and over	O	5	151	152	75	29	16	8	13	$24,200				

Section D: Marriage and Family Size

Item	Code	Officer or Wife	Distribution 0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
1a Ideal marriage age for average man	0=never 1=20 and under 2=21 and 22 6=29 and over	O	0	3	25	48	223	72	49	7	25.72	26.18	25.46		
		W	0	1	2	33	145	74	52	13	26.40	26.74	26.34		
1b Ideal marriage age for army officer	same as 1a	O	0	0	6	43	190	90	82	16	26.53				
		W	0	0	6	33	141	64	60	16	26.44				
1c Ideal marriage age for army air corps officer	same as 1a	O	4	0	9	41	168	92	81	32	26.67	27.15	26.69		
		W	0	0	4	34	133	67	54	28	26.45	26.88	26.38		
2a Ideal marriage age for average woman	same as 1a	O	0	87	187	96	46	3	0	8	22.03				
		W	0	44	131	86	49	4	0	6	22.40	22.58	22.18		
2b Ideal marriage age for army officer's wife	same as 1a	O	0	59	179	109	65	4	0	11	22.40				
		W	0	46	125	81	50	5	0	13	22.44				
2c Ideal marriage age for wife of army air corps officer	same as 1a	O	3	59	168	104	72	9	0	12	22.54				
		W	0	41	119	91	53	5	0	11	22.55	22.76	22.38		
3a Ideal number of children for average American family	0=0 1=1 6=6	O	0	0	79	190	123	21	4	10	3.24	2.00	4.00	2.99	3.62
		W	1	0	67	114	116	8	4	10	3.24	2.00	4.00	2.66	3.54
3b Ideal number of children for army officer's family	0=0 1=1 6=6	O	6	5	172	162	64	6	2	10	2.72				
		W	2	4	164	99	35	2	0	14	2.55				
3c Ideal number of children for family of army air corps officer	0=0 1=1 6=6	O	19	11	185	142	52	5	2	11	2.53			2.00	3.12
		W	5	13	170	83	37	0	0	12	2.44			1.87	2.89

Section D (*continued*)

Item	Code	Officer or Wife	Distribution 0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
4a Estimate of spouse's opinion on question 3a	0=0 1=1 5=5 or more 6=divorced or single	O	3	1	92	131	98	15	46	41	3.09				
		W	4	1	49	95	56	11	0	106	3.10				
4b Estimate of spouse's opinion on question 3c	0=0 1=1 5=5 or more 6=divorced or single	O	15	13	175	99	44	3	46	32	2.44				
		W	10	13	149	92	26	0	0	30	2.38				

Section E: Effect of Various Factors on Family Size

(Below are listed various changes in conditions which, it has been suggested, *might* influence the size of some families. Assume that the particular change has been effected at the time you originally entered the service. Assume also that all other factors except this particular one remained the same. Indicate the increase or decrease of the figure which you have shown in question *a* which this single change would have brought about.
Note: After answering Question *a*, 31 officers and 18 wives omitted the rest of the section.)

Item	Code	Officer or Wife	0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
a Total number of children have had plus expecting to have	0=0 1=1 6=6 or more	O	64	61	179	79	29	1	7	7	1.95				
		W	39	51	140	61	17	3	2	7	1.95				
1 If *retirement pay was equal* to full pay	0=decrease −3 3=no difference 6=increase +3	O	0	0	0	281	72	33	3	7	0.38				
		W	0	0	0	201	53	18	4	26	0.37				
2 If first substantial increase in pay was one year earlier	same as 1	O	0	0	0	333	47	2	1	8	0.14				
		W	0	0	0	237	33	4	1	27	0.16				
3 If educational expenses of your children were provided for by a fund set aside by a "rich uncle"	same as 1	O	0	0	0	176	118	73	22	7	0.55			0.91	0.86
		W	0	0	0	158	80	45	10	29	0.73			0.67	0.82

SECTION E (*continued*)

Item	Code	Officer or Wife	Distribution 0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
4 If wife could have children only by caesarian operation	same as 1	O	54	99	114	116	0	0	0	13	—1.22				
		W	9	46	71	139	5	0	2	30	—0.66				
5 If normal and healthy children were *guaranteed* by advances in genetic and medical theory	same as 1	O	0	0	0	322	36	26	5	7	0.26				
		W	0	0	0	224	31	10	7	30	0.25				
6 If all sons were to be granted admission to West Point or other service schools	same as 1	O	0	0	0	293	58	29	9	7	0.36				
		W	0	0	0	221	37	9	6	29	0.26				
7 *If* sons were guaranteed success in the occupation chosen as a result of new developments in educational and vocational counselling methods	*same* as 1	O	0	0	0	284	55	34	11	12	0.40				
		W	0	0	0	209	45	13	10	25	0.36				
8 If pay schedules in Air Corps were increased $50 a month	same as 1	O	0	0	0	233	130	23	2	8	0.47				
		W	0	0	0	176	78	15	5	28	0.44				
9 If purchasing power of dollar was reduced to thirty cents	same as 1	O	21	91	117	154	0	0	0	15	—0.93				
		W	18	48	66	138	0	0	0	32	—0.77				
10 If $1000 was provided at birth of each child by fund set up by a "rich uncle"	same as 1	O	0	0	0	273	83	25	4	11	0.38			0.41	0.36
		W	0	0	0	193	60	12	7	30	0.39			0.35	0.40

SECTION E (continued)

Item	Code	Officer or Wife	Distribution 0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
11 If a monthly income of $100 for wife and $50 per month for each child to the age of 21 was provided by government in case of death of an officer	same as 1	O	0	0	0	131	132	59	15	9	0.76			0.94	0.76
		W	0	0	0	154	73	29	14	27	0.65			0.79	0.55
12 If army pay schedules provided a 10% increase in base pay at birth of each child	same as 1	O	0	0	0	233	104	41	9	9	0.55				
		W	0	0	0	164	77	24	7	30	0.54				
13 If excellent nursery schools were provided at all posts	same as 1	O	0	0	0	358	20	6	2	10	0.10				
		W	0	0	0	251	13	7	2	29	0.12				
14 If person you married did not want children	same as 1	O	51	127	77	116	0	0	0	25	—1.29				
		W	29	76	66	95	0	0	0	38	—1.15				
15 If married life was less happy than anticipated	same as 1	O	48	121	71	133	0	0	0	23	—1.21				
		W	25	72	59	102	4	0	0	40	—1.05				
16 If maid service of high quality at low cost was assured	same as 1	O	0	0	0	341	31	11	3	10	0.15				
		W	0	0	0	219	43	12	4	24	0.28				
17 If predetermination of sex of children were made possible by medical science	same as 1	O	0	0	2	346	30	10	1	7	0.13				
		W	0	0	1	229	42	4	2	24	0.20				
18 If social activities in Air Corps were cut in half	same as 1	O	0	0	0	351	30	5	1	9	0.11				
		W	0	0	0	255	17	3	1	26	0.09				

SECTION E (*continued*)

Item	Code	Officer or Wife	Distribution 0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
19 If accidents in Army Air Corps were no more numerous than those of civil life	same as 1	O	0	0	0	288	68	24	9	7	0.37				
		W	0	0	0	192	56	22	6	26	0.42				
20 If person you married wanted a large family	same as 1	O	0	0	0	186	117	64	16	13	0.76				
		W	0	0	0	143	81	32	15	31	0.69				
21 If birth control were extremely simple and 100% effective	same as 1	O	0	8	9	356	7	7	1	8	0.00				
		W	2	3	9	245	8	4	2	29	0.00				
22 If person you married had exceptional ability to manage children well and to guide their personality development	same as 1	O	0	0	0	308	50	24	6	8	0.30				
		W	0	0	0	223	27	12	10	30	0.29				
23 If painless and safe childbirth were assured by advances in medical science	same as 1	O	0	0	0	277	68	28	15	8	0.45				
		W	0	0	0	198	44	19	8	33	0.39				
24 If flying hazards in Army Air Corps were doubled by more complex equipment	same as 1	O	20	81	79	205	0	0	0	11	—0.77				
		W	14	45	59	149	0	0	0	35	—0.68				

Section F: Marital Status and Family Plans

Part I, For Officers Who Have Never Been Married (There were 38 single officers.)

1	Are you engaged to be married	0=no 1=yes	O	28	9					1	.24
2	How soon do you expect to be married	0=within 6 months 1=within 1 year 2=within 2 years 3=within 5 years 4=perhaps sometime 5=probably never	O	4	2	4	8	12	7	1	3.16
5	Do you think you are too young to marry	0=no 1=yes	O	37	0					1	0
6	Do you think you are too old to marry	0=no 1=yes	O	38	0					0	0
7	Do you regret that you are not married	0=no 1=yes	O	23	14					1	.38
8	Have your parents been responsible in any way for the fact that you are not married	0=no 1=yes	O	27	11					0	.29
9	Do you regret that you did not marry at an early age	0=no 1=yes	O	30	8					0	.21
10	*Do you feel* that you have too many debts to marry	0=no 1=yes	O	35	3					0	.08
11	Do you feel that your income is too small to support a wife	0=no 1=yes	O	30	8					0	.21
12	Would you be married had your financial position been more secure in the past	0=no 1=yes	O	19	19					0	.50

SECTION F, Part I (*continued*)

Item	Code	Officer or Wife	Distribution 0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
14 Have you failed to find the right person	0=no 1=yes	O	16	22						0	.58				
15 Have you been disappointed in love	0=no 1=yes	O	30	8						0	.21				
17a First choice if you were to marry	0=beauty 1=brains 2=money 3=good disposition 4=good health 5=social prestige 6=good homemaker	O	1	8	1	11	13	0	4	0					
17b Second choice if you were to marry	same as 17a	O	3	10	0	11	10	1	3	0					

Section F, Part II, For All Officers Who Are or Have Been Married and All Wives (There were 389 married officers and 320 wives. Questions 1, 2, 7*abc*, and 18*a* include single officers.)

Item	Code	Officer or Wife	Distribution 0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
1 How many times have you been engaged to be married	0=never or "not engaged" 1=once 4=4 or more times	O	21	231	112	30	8			25	1.44			1.66	1.51
		W	0	133	122	37	20			8	1.82			2.02	1.72
2 Age when you first began having dates	0=14 or under 1=15 or 16 6=25 or over	O	67	181	105	29	9	5	2	29	16.25			16.44	16.31
		W	74	163	67	9	0	0	1	6	15.63			15.68	15.67

SECTION F, Part II (*continued*)

Item	Code	Officer or Wife	Distribution 0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
3c Have you remarried	0=*no* 1=yes	O W	8 0	39 55						0 9	.83 1.00			.93	.91
3d Cause of separation from first wife (or husband)	0=incompatibility 1=death 2=drinking 3=adultery 4=cruelty 5=spouse in love with another 6=desertion, non-support, or lack of responsibilty	O W	21 11	13 8	1 5	1 2	1 2	1 1	1 2	8 15					
5 Length of engagement	0=11 months and under 1=12 to 24 months 5=60 months or more 6=not engaged	O W	229 208	80 58	26 20	5 4	8 7	11 9	40	27 14					
7abc Discussion of plans for children prior to marriage	0=never discussed plans 1=plans discussed but not definite 2=made definite plans	O W	122 114	167 122	104 67					34 17					
7c Number of children planned for prior to marriage	0=0 1=1 6=6 or more	O W	14 3	6 1	36 29	32 21	15 10	0 0	1 5	0 0	2.31 2.72			1.06 2.10	2.95 3.39
8a Age at marriage	0=20 and under 1=21 and 22 6=31 and over	O W	1 74	59 92	92 72	96 47	81 18	34 5	46 10	0 2	26.40 22.83			27.32 23.85	26.54 22.70

Section F, Part III, For Married Officers (and Officers' Wives) Who Have Had No Children and Are Not Now Expecting Any (This includes 105 officers and 82 wives.)

1 Lacked money for the immediate expenses of having a child	0=no 1=yes	O W	62 51	39 28						4 3	.39 .35
2 Uncertain of the ability to meet the necessary later expenses of raising and educating the child	0=no 1=yes	O W	72 62	29 17						4 5	.29 .22
3 Support of dependents (father, mother, brother, etc.) has been a factor in your having no children (includes single officers' failure to marry because of this factor)	0=no 1=yes	O W	114 69	25 8						4 5	.18 .10
5 Has your wife been afraid of childbirth	0=no 1=yes	O W	87 70	12 8						6 4	.12 .10
6a Do you regret not having children	0=no 1=yes	O W	44 30	56 50						5 2	.56 .62
6b Does your spouse regret not having children	0=no 1=yes	O W	37 33	58 42						10 7	.61 .60
7 Number of children wanted	0=none 1=one child 5=five children or more 6=no number given or indifferent	O W	27 13	8 9	35 30	20 15	4 7	1 0	5 2	7 6	1.68 1.92
8a Efforts to have children	0=no efforts made 1=efforts unsuccessful	O W	47 38	52 37						6 7	.53 .49
8b If efforts unsuccessful, have you consulted a doctor	0=no 1=yes	O W	13 3	59 34						6 7	.75 .92

SECTION F, Part III (*continued*)

Item	Code	Officer or Wife	Distribution 0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
8c To what did the doctor attribute your difficulty	0=wife's health 1=organs of wife defective 2=organs of officer defective 3=doctor unable to diagnose	O W	5 3	14 11	4 2	14 15				2 3					
8d If not wishing children, what is your most important reason	0=financial 1=do not like or do not want children (either officer or wife) 2=wife's health or defective organs of wife 3=responsibility for sibs 4=too old 5=flying hazard 6=other	O W	5 1	9 2	5 2	1 0	1 5	5 2	3 4	15 10					
9a Does your spouse want any children	0=no 1=yes	O W	16 18	78 53						11 11	.83 .75				
9bc If your spouse does not want children, what is the most important reason from her or his point of view	0=fear of childbirth, wife's health, defective organs of wife, caesarian necessary 1=sterility 2=do not want or do not like children (either officer or wife) 3=financial 4=later in marriage 5=flying hazard 6=other reasons and unknown	O W	23 15	6 7	11 3	6 8	18 6	3 5	24 21	15 10					

SECTION F, Part III (*continued*)

Item	Code	Officer or Wife	Distribution 0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
10 Have you ever seriously considered adopting a child	0=no 1=yes	O W	70 49	29 29						6 4	.29 .37				
11 Number of children not your own which you are rearing	0=no children 1=one child 2=two children	O W	81 62	14 10	1 3					9 7					
11c Status of children not your own	0=wife's (or husband's) child adopted 1=other adopted 2=wife's child not adopted 3=relative's child not not adopted	O W	1 4	7 5	3 0	4 4				0 0					

Section F, Part V, For Married Officers (and Wives) Who Have Had Children, Living and Deceased (Questions 10*a*; 13*a*; 14*a*; 26*a*; 26*b*1, 2, 3, 4, 8, 9, and 10; and 27*a* include 12 officers and 11 wives who are now expecting their first child.)

Item	Code	Officer or Wife	Distribution 0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
2a Number of male children	0=first born 4=fifth born	O W	139 111	79 69	28 23	9 6	3 3			1 0	1.67 1.68				
2b Number of female children	0=first born 5=sixth born	O W	132 116	68 57	18 18	9 7	3 0	3 1		0 0	1.68 1.60				

SECTION F, Part V (*continued*)

Code	Item	Officer or Wife	Distribution 0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
10a Was the pregnancy of the first born intentional	0=definitely planned 1=accidental, i.e., conception occurred despite some attempted method of control 2=not definitely planned but no preventive method used	O	109	55	116					4					
		W	100	59	78					1					
10b Was the pregnancy of the second born intentional	same as 10a	O	52	37	51					3					
		W	50	36	35					1					
10c Was the pregnancy of the third born intentional	same as 10a	O	11	16	18					0					
		W	14	17	8					1					
10d Was the pregnancy of the forth, fifth, and sixth born intentional	same as 10a	O	4	12	10					1					
		W	2	8	7					0					
11a If the pregnancy of the first born was definitely planned, how many months after you stopped taking precautions did conception occur	0=0—1 month 1=2—3 months 5=10 months or more	O	52	16	10	7	3	12		8	1.9*				
		W	41	16	9	7	4	11		6	2.3*				
11b Time for conception for all other births after precautions were stopped	same as 11a	O	31	13	5	3	5	5		9	2.0*				
		W	28	9	6	5	3	12		7	2.5*				
12a How much fear did your wife have of childbirth in regard to your first born	0=little or none 1=some 2=much	O	158	89	21					4	0.49				
		W	158	50	17					2	0.37				

*These figures are medians rather than means.

Item	Code	Officer or Wife	Distribution 0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
12b How much fear did your wife have of childbirth in regard to your second born	same as 12a	O	89	40	11					4	0.44				
		W	82	29	10					2	0.40				
12c Wife's fear of childbirth—third born	same as 12a	O	33	10	1					0	0.27				
		W	26	9	3					1	0.39				
12d Wife's fear of childbirth—fourth, fifth and sixth born	same as 12a	O	18	7	1					1	0.35				
		W	11	5	1					0	0.41				
13a How did you feel when you first knew that your wife was pregnant for the first time (and with wife's feeling)	0=glad 1=indifferent 2=worried 3=sorry	O	201	21	40	19				3					
		W	167	13	29	27				2					
13b Your feeling when first knowing that your wife was pregnant for the second and third times	same as 13a	O	129	16	27	13				3					
		W	113	12	19	16				2					
13c Your feeling when first knowing that your wife was pregnant for the fourth, fifth and sixth times	same as 13a	O	10	8	4	2				3					
		W	6	5	2	4				0					
14a How did your spouse feel when she (or he) first knew that she was pregnant for the first time	same as 13a	O	201	7	51	21				4					
		W	170	22	28	14				4					
14b Spouse's feeling when knowing that she was pregnant for the second and third times	same as 13a	O	125	19	27	11				6					
		W	111	13	15	13				10					

Item	Code	Officer or Wife	0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
14c Spouse's feeling when knowing that she was pregnant for the fourth, fifth and sixth times	same as 15a 4=mad (fourth child)	O	9	4	6	2	1			5					
		W	8	2	2	3				2					
15 Were you disappointed at first in the sex of the child (all births combined)	0=male yes 1=male no 2=female yes 3=female no	O	5	249	64	165				9					
		W	21	136	33	167				4					
17a Did the care of your first born disrupt your previous normal mode of life	0=very much 2=some 4=practically not at all 6=child died at birth	O	22	55	103	44	54		14	0	2.05				
		W	20	39	78	35	46		10	1	2.21				
17b Did the care of your second, third, fourth, fifth and sixth born disrupt your previous normal mode of life	same as 17a	O	9	19	83	49	45		8	2					
		W	11	14	52	39	48		9	6					
20 How did you feel when your first child was born	0=very glad 2=indifferent 4=bitter (child died)	O	210	42	9	1	3			7	0.28				
		W	202	18	1	0	2			4	0.15				
21 How did your spouse feel when your first child was born	same as 20	O	224	34	3	1	5			7	0.21				
		W	186	23	10	3	1			4	0.25				
22 Amount of wife's worry about losing her figure after birth of first child	0=much 3=practically none	O	25	81	54	106				6	1.91			1.87	2.01
		W	21	42	45	117				2	2.15			1.83	2.37

SECTION F, Part V (*continued*)

Item	Code	Officer or Wife	Distribution 0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
23 How did the care of your first child influence your activities (or your husband's activities)	0=made no difference 1=stayed at home more 2=social activities curtailed 3=had to rebudget expenses 4=1 and 2; 1 and 3; and 2 and 3 5=1 and 2 and 3 6=child died	O W	41 96	99 50	21 11	14 22	48 28	39 13	5 2	7 5					
24 How did the care of this first child influence your wife's activities	0=made very little difference 1=hampered social activities 2=altered circle of friends 3=1 and 2; 1 and indoors more; and stayed home more 4=care of child 5=hastened divorce 6=child died	O W	104 110	132 87	11 6	10 16	2 0	0 1	3 2	10 5					
25 Effect of first child on size of family previously planned	0=desire larger family 1=desire smaller family 2=no difference	O W	79 83	37 47	118 64					38 53					
26a If there were more than 18 months between marriage and birth of first child, was this interval deliberately planned	0=no 1=yes	O W	56 32	86 81						15 15	.61 .72				

SECTION F, PART V (*continued*)

Item	Code	Officer or Wife	Distribution 0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
26 b1 Was this interval due to intervening miscarriages	0=no 1=yes	O W	109 84	20 10						25 31	.16 .11				
26 b2 Was this interval due to your living apart from your spouse	0=no 1=yes	O W	123 88	3 4						28 33	.02 .04				
26 b3 Was this interval due to a desire for a period of freedom from the responsibility of children	0=no 1=yes	O W	52 40	74 59						28 26	.59 .60	.55 .50	.50 .64	.53 .40	.51 .64
26 b4 Interval due to delay in establishing satisfactory sexual relationship	0=no 1=yes	O W	117 84	10 5						27 37	.08 .06				
26 b5 Interval due to fear of possible hereditary defects	0=no 1=yes	O W	119 86	1 0						27 32	.01 .00				
26 b6 Interval due to change of residence	0=no 1=yes	O W	109 75	12 16						26 27	.10 .18				
26 b7 Interval due to wife's fear of childbirth	0=no 1=yes	O W	113 79	6 8						28 51	.05 .09				
26 b8 Interval due to lack of money for the immediate expense of having a child	0=no 1=yes	O W	72 61	58 36						24 28	.45 .37	.40 .32	.61 .33	.47 .26	.38 .29
26 b9 Interval due to uncertainty of ability to meet the necessary later expenses of raising and educating the child	0=no 1=yes	O W	95 71	33 24						26 30	.26 .25	.32 .15	.37 .26	.29 .25	.15 .14

Item	Code	Officer or Wife	Distribution 0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
26 b10 Interval due to desire to be certain of compatibility	0=no 1=yes	O	95	31						28	.25	.35	.23	.23	.22
		W	62	38						25	.38	.38	.29	.33	.29
26 b11 Other reasons for interval	0=defective reproductive organs 1=wife's health 2=frequent moving 3=war and flying hazard 4=to be near good doctor and fear of caesarian 5=financial 6=reason unknown	O	4	7	1	3	1	1	3	31					
		W	5	6	0	3	1	0	0	21					
27a Number of additional children planned	0=none 1=one child 6=some (number unspecified)	O	164	69	27	5	0	0	1	18	.52				
		W	157	67	20	3	0	1	1	9	.53				
27b If no more children planned, is physical disability or sterility a cause	0=no 1=physical disability 2=operation to officer 3=operation to wife	O	144	7	2	4				21					
		W	124	8	2	3				8					
27c Chief reason for planning no more children (includes multiple answers and reasons for limiting family even though planning at least one more child)	0=financial and cost of insurance 1=age and age of present children 2=consideration for wife's health 3=army living conditions 4=more children would be inconvenient 5=satisfied with present size of family 6=other reasons	O	91	36	46	23	14	16	17	24					
		W	76	35	41	28	9	15	13	10					

Section G: Reasons for Having Children

(The following have been reported by various persons as reasons why parents have children. Rate your belief regarding the importance to you of each of these reasons for having children.) N.B. The means for prolific and non-prolific group in this section are based on 53 and 68 officers and 51 and 65 wives respectively.

Item	Code	Officer or Wife	Distribution 0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
1 Importance to you of having children in order to carry on the family name and social traditions	0=none 1=little 2=some 3=much 4=great	O	87	75	136	67	49			13	1.80			1.62	2.18
		W	95	46	94	34	37			14	1.58			1.37	1.43
2 Importance to you of having children in order to have companionship of young children	same as 1	O	20	26	83	152	133			13	2.85			2.53	3.29
		W	15	17	57	100	118			13	2.94			2.51	3.03
3 Importance in carrying on the family enterprises	same as 1	O	266	76	47	17	4			17	.58			0.57	0.51
		W	211	43	37	4	4			21	.48			0.29	0.42
4 Importance in carrying on the professional traditions of the family	same as 1	O	250	65	76	31	9			16	.84			0.83	0.75
		W	196	58	45	12	8			25	.65			0.47	0.42
5 Importance in providing security for old age	same as 1	O	351	44	13	2	0			17	.19			0.15	0.25
		W	265	23	11	2	1			18	.18			0.10	0.18
6 Importance in providing a feeling of contentment in old age	same as 1	O	76	42	127	110	56			16	2.07			1.66	2.50
		W	46	23	72	83	83			13	2.44			1.84	2.37
7 Importance in perpetuating the social class of which the parents are members	same as 1	O	142	69	112	55	35			14	1.45			1.49	1.60
		W	139	43	68	32	23			15	1.20			0.94	1.25
8 Having children is considered a social obligation	same as 1	O	209	65	97	25	15			16	.96			1.13	1.25
		W	134	45	46	19	13			15	.80			0.65	1.05

Item	Code	Officer or Wife	Distribution 0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
9 Having children is considered a religious obligation	same as 1	O	315	48	54	8	9			13	.45			0.36	0.53
		W	219	31	28	11	14			17	.58			0.47	0.77
10 A family is not complete without children	same as 1	O	24	22	56	137	178			10	3.01			2.60	3.41
		W	15	6	25	55	210			9	3.41			2.84	3.54
11 Children provide one of the most satisfying outlets for self-expression	same as 1	O	109	57	114	78	54			15	1.78			1.68	2.25
		W	58	42	77	62	64			17	2.11			2.00	2.00
12 Children are needed to carry on the traditions of the country	same as 1	O	84	64	116	77	71			15	1.97			1.98	2.10
		W	75	37	62	62	71			13	2.06			1.94	1.97
13 Children tend to make the family more stable	same as 1	O	31	20	93	139	131			13	2.77			2.53	3.03
		W	28	21	55	98	108			10	2.76			2.20	2.97
14 Children assist the parents in maintaining a youthful point of view	same as 1	O	56	46	112	126	76			11	2.29			2.00	2.57
		W	25	23	65	94	105			12	2.75			2.37	2.74
15 Children provide a basis for social contacts	same as 1	O	220	85	81	19	6			16	.80			0.81	0.85
		W	133	48	81	22	23			13	1.20			1.02	1.14
16 The stabilizing influence of children is recognized by occupational superiors	same as 1	O	212	69	72	40	17			17	.98			0.83	0.81
		W	134	42	74	25	21			24	1.18			1.35	0.78
17 Children tend to prevent separation and divorce	same as 1	O	111	47	124	78	54			13	1.80			1.36	1.94
		W	78	54	79	58	39			12	1.76			1.57	1.71
18 The physical and mental health of the mother is improved by having children	same as 1	O	159	70	105	49	20			24	1.26			0.85	1.65
		W	120	52	73	48	13			14	1.29			0.88	1.46

Section G (*continued*)

Item	Code	Officer or Wife	0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
19 The obligations of parenthood tend to compel the parents to "settle down"	same as 1	O	133	73	118	63	26			14	1.46			1.17	1.41
		W	105	45	86	47	22			15	1.46			1.02	1.40
20 Children offer an *opportunity* for tactful release from too much social activity	same as 1	O	209	83	87	25	8			15	.88			0.75	1.18
		W	154	57	53	28	14			14	.99			0.61	1.00
21 Children provide *the perfect* fusion of the personalities of two individuals who are in love	same as 1	O	75	46	96	100	96			14	2.25			2.17	2.47
		W	41	21	61	68	113			16	2.63			2.31	2.77
22 One of the *purposes* of marriage is to produce children	same as 1	O	108	47	119	89	52			12	1.83			1.25	1.90
		W	54	38	67	72	77			12	2.26			1.69	2.43
23 The producing and *rearing of a well-rounded* individual typifies the highest form of artistic expression	same as 1	O	135	42	95	79	61			15	1.73			1.49	1.87
		W	67	24	66	59	83			21	2.22			1.94	2.12
24 A woman's life is not complete without the experience of childbirth	same as 1	O	143	54	94	68	54			14	1.60			1.49	1.79
		W	68	35	45	56	99			17	2.27			1.76	2.29
25 Children "*soften*" the personalities of the parents	same as 1	O	117	73	120	68	53			16	1.58			1.32	1.62
		W	75	42	69	60	54			20	1.92			1.35	1.82
26 The rearing of children gives a woman's personality a *more* feminine touch	same as 1	O	135	59	112	72	31			18	1.52			1.28	1.71
		W	94	55	78	44	53			16	1.56			1.10	1.45

SECTION G (*continued*)

	Item	Code	Officer or Wife	Distribution 0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
27	Parents have children because their friends have children	same as 1	O	324	53	30	4	0			16	.30			0.23	0.32
			W	253	24	22	2	4			15	.30			0.33	0.28
28	The parents enjoyed being children themselves	same as 1	O	163	79	86	58	23			18	1.26			0.98	1.65
			W	103	59	66	49	30			13	1.49			1.37	1.28
29	Children provide a type of physical immortality	same as 1	O	195	75	80	35	20			22	1.04			0.92	1.18
			W	122	35	80	40	25			18	1.37			1.12	1.37
30	Children provide excellent recreation	same as 1	O	93	51	133	104	31			15	1.83			1.66	2.09
			W	55	43	79	74	54			15	2.10			2.02	2.32
31	Raising children is an extremely interesting hobby	same as 1	O	115	57	107	92	38			18	1.71			1.25	2.06
			W	55	29	60	69	95			12	2.39			2.27	2.49
32	Children are very helpful around the house	same as 1	O	238	104	59	6	4			16	.62			0.55	0.57
			W	153	71	55	16	11			14	.89			0.84	0.92
33	Watching children grow up is a lot of fun	same as 1	O	38	38	111	147	79			14	2.46			2.13	2.84
			W	23	13	48	107	118			11	2.92			2.75	2.98
34	Good stock ought to reproduce itself	same as 1	O	60	42	91	109	112			15	2.41			2.32	2.74
			W	41	18	65	72	110			14	2.65			2.45	2.65
35	Being married to a person who wants children	same as 1	O	43	26	117	121	107			13	2.54			2.11	2.57
			W	17	14	64	93	118			14	2.92			2.57	2.78
36	To provide company in old age	same as 1	O	110	66	125	64	46			16	1.68			1.40	2.03
			W	81	49	63	56	57			14	1.87			1.27	1.94
37	Having children around is a lot of fun	same as 1	O	46	45	117	133	71			15	2.33			1.89	2.74
			W	27	18	59	95	108			13	2.78			2.41	3.02
38	To provide company for the wife when the husband is away	same as 1	O	99	65	135	73	40			15	1.73			1.53	1.81
			W	75	44	71	62	53			15	1.91			1.69	1.88

ALLPORT-VERNON: A STUDY OF VALUES

Scale	Code	Officer or Wife	0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
A1 Theoretical	0=0—9 1=10—19 5=50—59	O	0	15	135	231	36	0		4	31.44	32.29	30.56	31.39	30.37
		W	2	58	163	73	8	0		3	25.40	24.76	25.52	25.34	25.59
A2 Economic	same as A1	O	0	4	82	217	113	1		4	35.12	36.50	34.58	35.14	35.05
		W	0	15	120	144	25	0		3	30.56	31.62	29.69	31.05	29.86
A3 Aesthetic	same as A1	O	6	125	199	76	11	0		4	23.45	24.33	22.92	24.57	22.11
		W	0	4	117	155	45	3		3	31.91	31.97	31.65	33.08	31.28
A4 Social	same as A1	O	0	9	226	172	10	0		4	28.91	28.04	29.89	28.68	30.04
		W	0	9	117	159	19	0		3	30.43	30.24	30.70	30.95	30.28
A5 Political	same as A1	O	0	0	32	174	193	18		4	39.19	40.29	39.19	40.64	38.87
		W	0	6	72	175	51	0		3	33.60	34.86	32.64	34.29	32.62
A6 Religious	same as A1	O	14	171	155	63	13	1		4	21.89	18.67	22.89	19.57	23.65
		W	3	39	138	90	32	2		3	27.98	26.57	29.80	25.15	30.14

BERNREUTER *Personality Inventory*

Scale	Code	Officer or Wife	0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
B1-N A measure of neurotic tendency (High scores indicate emotional instability)	0= —161 to —240 1= —81 to —160 6=240 to 319	O	71	186	115	35	10	2	0	8	—90.7			—89.3	—83.9
		W	10	78	108	74	38	4	0	8	—22.7			—28.1	—10.9
B2-S A measure of self-sufficiency (High scores indicate self reliance)	same as B1-N	O	0	1	56	250	111	1	0	8	50.7	57.5	48.3	56.5	46.7
		W	0	38	131	114	29	0	0	8	—5.1	—5.9	—5.5	9.1	—9.3
B3-I A measure of introversion (High scores indicate introversion)	same as B1-N	O	0	97	263	55	4	0	0	8	—46.9			—45.7	—43.3
		W	0	22	159	120	11	0	0	8	—10.3			—13.5	—3.3

BERNRUETER *Personality Inventory (continued)*

Scale	Code	Officer or Wife	Distribution 0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
B4-D A measure of dominance-submission (High scores indicate dominance)	same as B1-N	O	0	3	29	175	199	15	0	8	74.9			73.3	74.1
		W	0	20	94	141	55	4	0	8	21.9			25.1	13.7
F1-C A measure of confidence in oneself (High scores indicate feelings of inferiority)	same as B1-N	O	40	153	139	67	16	5	1	8	—65.1	—64.1	—63.7	—62.5	—56.5
		W	4	51	85	96	53	22	3	3	17.5	12.3	27.1	7.9	31.5
F2-S A measure of sociability (High scores indicate a non-social, solitary, or independent type)	same as B1-N	O	1	21	163	197	37	0	0	8	5.5	17.5	3.9	14.9	5.5
		W	3	37	164	97	11	0	0	8	—21.5	—22.9	—17.1	—7.5	—22.1
Attitude Report Form															
Optimism-Pessimism Scale (High scores indicate optimism)	0=48—53 1=54—59 6=84—89	O	1	16	100	163	105	22	4	16	68.78	67.09	69.67	67.36	69.72
		W	1	17	82	115	57	13	2	33	67.99	68.15	68.00	65.97	68.08

Item	Code	Officer or Wife	Distribution 0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
Officer's Rating of the Extent of His Satisfaction with Present Occupation															
Satisfaction	0=extremely well satisfied 2=moderately well satisfied 5=not at all satisfied	O	176	195	34	2	0	1		15	0.67			0.54	0.73
Officer's Rating of the of the Extent to Which an Individual of His Type Would Be Successful As an Air Corps Officer															
Success	0=extremely successful 3=neither successful nor unsuccessful 6=extremely unsuccessful	O	66	212	122	3	0	1	1	18	1.18			1.18	1.18
Strong *Vocational Interest Blank* (Officers)															
1 Engineer	0=A 1=B+ 2=B 3=B− 4=C	O	173	89	79	14	63			5	1.29			1.41	1.29
2 Physician	same as 1	O	37	61	94	46	180			5	2.65			2.65	2.66
3 Lawyer	same as 1	O	19	61	79	92	167			5	2.78			2.59	2.91
4 *Personnel* manager	same as 1	O	37	83	120	42	136			5	2.38			2.37	2.41
5 Purchasing agent	same as 1	O	31	87	145	56	101			5	2.26			2.05	2.35
6 Architect	same as 1	O	5	14	55	50	294			5	3.47			3.49	3.44
7 Journalist	same as 1	O	24	53	59	66	216			5	2.95			2.71	3.06

STRONG *Vocational Interest Blank* (OFFICERS) (*continued*)

	Item	Code	Officer or Wife	Distribution 0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
8	YMCA physical director	same as 1	O	16	44	71	45	242			5	3.08			3.13	2.94
9	Mathematician	same as 1	O	4	17	51	44	302			5	3.49			3.69	3.47

STRONG *Vocational Interest Blank* (WIVES)

	Item	Code	Officer or Wife	0	1	2	3	4	5	6	X	Mean	Mean ideal 2	Mean ideal 4	Mean planned family 0 or 1	Mean planned family 3 or more
1	Nurse	0=A 1=B+ 2=B 3=B— 4=C+ 5=C	W	68	58	54	47	49	33		7	2.16			2.55	2.01
2	Librarian	same as 1	W	24	22	58	64	64	77		7	3.14			2.86	3.25
3	Teacher in general	same as 1	W	10	22	25	27	39	188		7	4.03			4.09	4.16
4	Artist	same as 1	W	67	53	47	47	40	55		7	2.54			2.16	2.46
5	Author	same as 1	W	47	37	52	56	42	75		7	2.76			2.56	2.84
6	Housewife	same as 1	W	98	69	71	37	19	15		7	1.55			1.78	1.46
7	Life insurance sales	same as 1	W	2	6	11	30	41	219		7	4.46			4.34	4.59
8	Social worker	same as 1	W	5	15	21	49	54	165		7	4.05			4.17	4.02
9	Secretary-Stenographer	same as 1	W	105	67	63	40	27	7		7	1.48			1.56	1.57

REFERENCES

1. ALLPORT, G. W., & VERNON, P. A Study of Values. Boston: Houghton-Mifflin, 1931.
2. BERNREUTER, R. G. The Personality Inventory. Stanford University: Stanford Univ. Press, 1931.
3. DAVIS, KATHERINE B. Factors in the Sex Life of Twenty-Two Hundred Women. New York: Harpers, 1929.
4. KELLY, E. L. Unpublished Schedules Developed for a Longitudinal Study of Marriage.
5. LORIMER, F., & OSBORN, F. Dynamics of Population. New York: Macmillan, 1934.
6. NOTESTEIN, F. W., & KISER, C. V. Fertility of the social classes in the native white population of Columbus and Syracuse. *Human Biol.*, 1934, **6**, 595-611.
7. OSBORN, F. Preface to Eugenics. New York: Harpers, 1940. Pp. 312.
8. STRONG, E. K. Vocational Interest Blank for Men. Stanford University: Stanford Univ. Press, 1927.
9. ———. Vocational Interest Blank for Women. Stanford University: Stanford Univ. Press, 1933.
10. TERMAN, L. M. Psychological Factors in Marital Happiness. New York: McGraw-Hill, 1938.

Published as a separate and in *Genetic Psychology Monographs*, 1942, 26, 101-155.

A GENETIC STUDY OF GEOMETRICAL-OPTICAL ILLUSIONS*

SISTER ANNETTE WALTERS

Institute of Child Welfare, University of Minnesota

*Accepted for publication by John E. Anderson of the Editorial Board, received in the Editorial Office on August 11, 1941, and published at Provincetown, Massachusetts.

ACKNOWLEDGMENTS

The author acknowledges with gratitude the assistance and coöperation of the following persons:

Dr. Miles A. Tinker, who directed the study, and who has been unfailing in his helpfulness at every stage of the work.

Dr. John E. Anderson, whose constructive suggestions and illuminating criticisms have yielded considerable insight into problems of growth.

Dr. Palmer O. Johnson and Dr. Paul L. Dressel, who have been most generous in assisting with the statistical treatment of the data.

The school principals who provided subjects, and who arranged school programs so as to facilitate this research. The author is particularly indebted to:

Brother John Berchmans, Principal of Cretin Military Academy;
Sister Mary John, Principal of St. Joseph's Academy;
Sister Marie Ursule, Principal of Derham Hall High School;
Sister St. Hugh, Directress of St. Luke's School;
Sister Cecile, Directress of Nativity School;
Sister Emile, Directress of St. Mark's School;

Sister Antonius and Sister Maris Stella, who read the manuscript and made suggestions in regard to form.

The superiors of my community, the Sisters of St. Joseph of Carondelet, who generously arranged time schedules and provided clerical assistance to facilitate the research. Without their encouragement this study would not have been completed.

SISTER ANNETTE WALTERS

The College of St. Catherine
St. Paul, Minnesota

I. HISTORY OF THE PROBLEM

Although intensive research on optical illusions was not undertaken until the latter part of the nineteenth century, the problem is as old as psychology itself. Oppel's (31) study, in 1854, is reputed to have been the first experimental attack on the problem. It was this study which drew attention to the fact that there is a tendency to overestimate the length of vertical lines in relation to horizontal lines. Müller-Lyer's (4) publication of his *Optische Urteiltauschungen,* in 1880, stimulated interest in the illusion of the arrowhead and featherhead which bears his name. It is with the developmental aspect of these two illusions that the present research is concerned.

A. Explanatory Theories of Illusion

The universality of geometrical-optical illusions suggests that some factor basic to visual perception in general is at work. Hence, it is not surprising to discover that the theories which have been proposed to explain these illusions parallel the theories of visual perception itself. Proposed explanations may be conveniently grouped into two main categories, namely: (*a*) those which attribute illusions to peripheral factors, and (*b*) those which attribute them to central factors. Since the more important of both groups of these theories have been reviewed elsewhere (31, 92), no discussion of them will be included in this monograph.

B. Age and Susceptibility to Illusions

Experimental studies of age and susceptibility to illusions have yielded contradictory and ambiguous results. Binet (5), Van Biervliet (87), Rivers (66), and Pintner and Anderson (65) report that the Müller-Lyer illusion decreases with advance in age. Cramaussel (13), on the other hand, claims that this illusion increases with age. Hartman and Triche (25) take issue with both of the above points of view and hold that the illusions of children, when considered as a group, are not different from those of adults. None of the above studies, however, can be considered reliable because of the limited number of subjects used, the lack of adequate controls, and the failure to apply appropriate statistical techniques. This holds true also for studies of the vertical-horizontal illusion, such as that of Seashore and Williams (69), in which the illusion was found to be slightly less for adults than for children.

C. Culture and Susceptibility to Illusions

Rivers' (66) study appears to be the only one which deals with the relation between culture and susceptibility to illusion. Rivers found that the Toda and the Papuan are markedly more susceptible to the vertical-horizontal illusion than is the Englishman. The Englishman, on the other hand, was found to be more susceptible to the Müller-Lyer illusion.

D. Illusions in Animals

Warden and Baar (89) trained two ring doves to discriminate between a short (positive) and a long (negative) line placed at the end of a Yerkes-Watson discrimination apparatus. When the Müller-Lyer illusion figures were substituted for the lines the birds responded to the line which was apparently shorter. Similarly, Winslow (92) was able to demonstrate that chicks are susceptible to both the vertical-horizontal and to the Müller-Lyer illusions.

E. Intelligence and Susceptibility to Illusions

Studies by Williams (90) and by Crosland, Taylor, and Newsom (14, 15) lead to the conclusion that intelligence and susceptibility to illusion are not related. The results of the latter's studies, however, suggest that there is a negative relationship between high intelligence and a high degree of illusion.

F. Effects of Practice on Illusions

Judd's findings, in common with those of other investigators, show that the Müller-Lyer illusion disappears with practice. Lewis (47), however, claims that the illusion disappears with practice only when the figures are exposed for prolonged periods. When the figures are exposed only momentarily the illusion does not disappear. Seashore *et al.* (68) present evidence that the influence of practice depends upon a number of factors, such as attitudes and knowledge of results. Crosland, Taylor, and Newsom (14) find that practice results in an initial increase in Müller-Lyer illusion, followed by a gradual decrease. Evidence regarding the influence of practice on the vertical-horizontal illusion is less clear cut. Williams (90) reports that the illusion of the vertical does not decrease with practice. Valentine (86) reports that for at least three of his subjects the vertical-horizontal illusion was increased by practice.

G. Sex Differences

Williams' (90) study is the only one in which an analysis of sex differences was made. He found no significant differences between the records of girls and boys.

H. Development of Perception in General

Vernon (88) refers from time to time to the fact, so often stressed by Gestalt psychologists, that the perceptions of young children may differ considerably from those of adults, and may resemble more closely those of animals, especially of the primates. She holds, in common with the Gestalt school, that children are particularly likely to perceive the field as a series of undifferentiated meaningful wholes. If the Gestalt explanation of the Müller-Lyer illusion is correct, then one ought to find a progressive decrease in the amount of illusion with advancing age until perceptual maturity is reached. The present research is designed to throw light on this problem.

I. Summary and Prospectus

The foregoing survey, revealing as it does a diversity of conflicting evidence, indicates the need for more adequate information concerning the problem of geometrical optical illusions. As early as 1902, Judd (40) suggested that this might be accomplished by a thoroughgoing genetic study of the qualitative and quantitative aspects of the perception of such illusions. Several such studies have indeed been reported, but, in general, confidence cannot be placed in the results because of the limited number of subjects used and because of the lack of adequate controls. It has been demonstrated that a number of factors associated with the experimental set-up may influence the amount of illusion, and since these factors have varied from experiment to experiment, there is no way of generalizing from one study to another. Binet (5), for example, has shown in his study of the Müller-Lyer illusion that such factors as the order in which one compares the lines and the size of the figures used have a marked influence upon perception. Pieron (64) has also demonstrated that the angle of inclination of the obliques and the proportionate length of the obliques are important influences.

An adequate genetic study of geometrical-optical illusions would necessitate the use of an identical set-up for a comparable sampling of individuals at all age levels. It would involve the careful control

of all those conditions, other than age differences, which might in any way influence the results. It would require a technique for measuring illusions reliably from one age level to the next, and would make provision for tracing the development of the same group of children over a period of several years. The extent to which susceptibility to illusion is related not only to age, but also to such factors as general intelligence and sex would be an interesting phase of the investigation. Such a study ought to throw light not only upon the problem of perceptual development, but upon the broader subject of mental development as well.

II. THE PROBLEM AND THE METHOD

A. The Problem

The chief purpose of this investigation is to determine if susceptibility to certain geometrical-optical illusions varies significantly and predictably with age, and to analyze the growth trends in the perception of illusions. The figures selected for the study were the Müller-Lyer and the vertical-horizontal illusion, because the general theories of illusion which have been proposed all bear on one or the other of these two figures, and because generalizations from certain studies of these illusions are frequently quoted in discussions of perceptual development.

An analysis of the previous studies of this problem demonstrated their inadequacy from the point of view of methodology, the number of subjects, and the age range of subjects. Such generalizations as have been drawn from these studies are, therefore, suspect. In view of the renewed interest in perceptual development, as a result of recent trends in theoretical psychology, it appears highly desirable to have conclusive evidence regarding the genetic aspects of geometrical-optical illusions.

B. Experimental Set-up

1. *Apparatus and Test Conditions*

The Müller-Lyer figure was drawn in black on white cardboard and glued to a metal supporting frame which stood perpendicular to the table. The variable portion of the figure, glued to a metal slider, moved freely in a groove when the experimenter turned a knob attached to the back of the frame and slider. A rotary adjustment movement of this knob resulted in a linear motion of the slider. The variable line, having oblique lines turned outward, was 50 mm. at its minimum and 160 mm. at its maximum length. All oblique lines were 4 cm. long and were placed at 45° angles from the comparison lines. Small white paper circles were placed at the heads of the angles as terminal reference points in order to increase accuracy. The width of all lines was 1/16 inch. On the back of the slider was a scale from which the experimenter could determine the amount of illusion in terms of millimeters. A large grey cardboard screen hid all parts of the apparatus except the illusion board, and concealed the movements of the experimenter.

The vertical-horizontal figure was drawn in black on white card-

board and glued to a metal frame and slider as was the Müller-Lyer figure. The horizontal line, 128 mm. in length, served as the standard, and the vertical line, varying in length from 54 mm. to 180 mm., served as the variable. The slider was adjusted from the back by the experimenter, as with the Müller-Lyer figure, and a grey screen concealed the adjusting mechanism and the movements of the investigator. A scale, drawn on the back of the apparatus, revealed the amount of illusion in terms of millimeters.

The subjects were studied individually and under identical conditions from one age level to the next. Each observer sat with his back to the window (northern exposure) and placed his chin in a chin rest clamped to the table. The base of the illusion apparatus was placed on the table 40 inches away from the chin rest. Measurements were taken only on bright days and during the school hours. All of the data were collected personally by the writer.

Four trials with each illusion board were given in an *ABBA* order at one sitting. In the first trial (*A*) the variable line was placed at its minimum length and gradually made longer, whereas in the second trial (*B*) the variable line was placed at its maximum length and gradually made shorter. The variable line was moved at a rate of one millimeter per .66 second. The directions given to the subjects were as follows:

Directions for Müller-Lyer Experiment

Notice the line extending from this dot to this dot (experimenter points to white dots at boundaries of standard line) and the line extending from this dot to this dot (experimenter points to dots at boundaries of variable line). Now I am going to move this line (experimenter points to variable line) and I want you to tell me when it appears to be the same length as this line (experimenter points to standard line). Tell me when they look the same length. Watch carefully!

Directions for Vertical-Horizontal Experiment

In this experiment I am going to move the vertical line (experimenter points to vertical line) and I want you to tell me when it appears to be the same length as the horizontal line (experimenter points to horizontal line). Watch carefully!

In giving directions to young children it seemed clearer at times to refer to the vertical line as "the line that goes up and down" and to the horizontal line as "the line that goes from side to side." Sometimes, too, the child was asked to tell which of the two lines

was the "bigger" instead of the "longer," when the former word seemed to be better understood than the latter.

C. Reliability of Illusion Scores

As soon as a sufficient amount of data had been collected the reliability of the method was determined for several age groups. This was found to be sufficiently high to warrant continuing with the procedure. Moreover, a reference to the study of Crosland *et al.* (5) suggested that the reliability of such measures was not appreciably increased when the average scores were based upon six trials instead of four. Therefore, four was considered to be an adequate number of trials.

Split-half reliability coefficients obtained by correlating average scores on the first and second trials with average scores on the third and fourth trials, and corrected by the Spearman-Brown formula, are shown in Table 1 and Table 2. It will be seen that with the excep-

TABLE 1

Split-Half Reliability Coefficients for Müller-Lyer Figure Corrected by Spearman-Brown Formula

Age group	N	r	r_x
5^{0}- 6^{5}	110	.1922	.3224
6^{0}- 7^{5}	106	.5697	.7259
7^{0}- 8^{5}	124	.5812	.7351
8^{0}- 9^{5}	101	.7158	.8343
9^{0}-10^{5}	137	.6519	.7892
10^{0}-11^{5}	162	.7836	.8786
11^{0}-12^{5}	166	.7413	.8514
12^{0}-13^{5}	139	.7679	.8687
13^{0}-14^{5}	122	.7921	.8839
14^{0}-15^{5}	108	.6700	.8023
15^{0}-16^{5}	105	.9516	.9751
16^{0}-17^{5}	111	.8233	.9030
17^{0}-18^{5}	110	.8064	.8922
18^{0}-19^{5}	93	.9152	.9557

tion of Müller-Lyer scores for age groups 5^{0}-6^{5}, they are sufficiently high to warrant the making of group comparisons.

D. Selection of Subjects

The subjects for the experiment were selected with a view toward obtaining a normal and comparable distribution of intelligence at all age levels. At the lower ages they were pupils of two parochial

TABLE 2

SPLIT-HALF RELIABILITY COEFFICIENTS FOR VERTICAL-HORIZONTAL FIGURE CORRECTED BY SPEARMAN-BROWN FORMULA

Age group	N	r	r_{B}
5^{0}- 6^{5}	110	.7565	.8613
6^{0}- 7^{5}	106	.8353	.9102
7^{0}- 8^{5}	124	.8628	.9263
8^{0}- 9^{5}	101	.7690	.8694
9^{0}-10^{5}	137	.8420	.9142
10^{0}-11^{5}	162	.8305	.9074
11^{0}-12^{5}	166	.7926	.8843
12^{0}-13^{5}	139	.7627	.8633
13^{0}-14^{5}	122	.6765	.8070
14^{0}-15^{5}	108	.8236	.9032
15^{0}-16^{5}	105	.8136	.8972
16^{0}-17^{5}	111	.8868	.9400
17^{0}-18^{5}	110	.7926	.8843
18^{0}-19^{5}	93	.7989	.8882

schools, one of which was in an average city neighborhood, the other of which was in a somewhat superior neighborhood. At the upper age levels they were students of a Catholic high school for girls, a Catholic military academy for boys, and of a Catholic college for women. The majority of the college students were selected from the lower third of the freshman class, and, for the remainder, somewhat more were taken from the middle third than from the upper third of the group.

The Kuhlmann-Anderson *Intelligence Tests* for ages six to maturity were given to all subjects, since these tests constitute a continuous scale with tests of increasing difficulty from the lower to the upper age levels. Although there is reason to doubt the adequacy of this scale for the college students, the personal constants appear to be sufficiently comparable for age groups seven through 14 to justify their use in this experiment.

The Heinis *Personal Constant* was adopted as the measure of intelligence because it appears to be more constant than the *IQ* (10, 29, 32, 44) and permits greater accuracy of prediction. This would be an important consideration if test records secured sometime previous to the actual experiment were to be used. It was not necessary to do this, however, since the investigator was able to administer Kuhlmann-Anderson tests to all subjects within two months of the illusion experiment.

The number of subjects and the means and standard deviations of personal constants at each age level for which they were obtained are shown in Table 3.

TABLE 3

DISTRIBUTION OF PERSONAL CONSTANTS AT DIFFERENT AGE LEVELS

Age group	N	Mean PC	σ_{PC}
5^0- 6^5	110		
6^0- 7^5	106	104.23	4.16
7^0- 8^5	124	102.80	5.02
8^0- 9^5	101	104.43	4.91
9^0-10^5	137	104.86	4.86
10^0-11^5	162	103.45	4.82
11^0-12^5	166	103.53	4.97
12^0-13^5	139	102.56	4.86
13^0-14^5	122	102.19	4.99
14^0-15^5	108		
15^0-16^5	104		
16^0-17^5	111		
17^0-18^5	110		
18^0-19^5	93		

The age levels and the number of subjects used in the longitudinal study are shown in Table 4. The amount of time which elapsed

TABLE 4

AGE AND NUMBER OF SUBJECTS STUDIED OVER A THREE-YEAR PERIOD

Age at first testing	Age at second testing	N
5^0-6^5	8^0- 9^5	18
6^0-7^5	9^0-10^5	70
7^0-8^5	10^0-11^5	83
8^0-9^5	11^0-12^5	63

between the first and the second testing was from three to three and one-third years in every case.

III. THE GROWTH PHENOMENA

A. Mean Differences in Illusion Scores at Different Ages

1. *Müller-Lyer Illusion*

The means and standard deviations of Müller-Lyer illusion scores for ages six through 19 are given in Table 5 and are graphically

TABLE 5

Means and Standard Deviations of Illusion Scores at Different Age Levels

Age group	N	Müller-Lyer Illusion Mean	Müller-Lyer Illusion σ	Vertical-Horizontal Illusion Mean	Vertical-Horizontal Illusion σ
5^6-6^5	110	29.10	9.78	32.96	6.32
6^6-7^5	106	26.62	6.26	32.34	8.82
7^6-8^5	124	24.93	6.38	33.18	9.26
8^6-9^5	101	23.91	5.72	33.20	6.54
9^6-10^5	137	23.49	5.81	32.12	8.21
10^6-11^5	162	23.76	6.10	23.45	8.60
11^6-12^5	166	24.34	6.10	28.54	7.41
12^6-13^5	139	23.36	5.59	28.27	6.62
13^6-14^5	122	24.84	5.60	26.94	6.93
14^6-15^5	108	26.71	2.34	24.96	4.11
15^6-16^5	104	25.51	2.94	23.77	4.01
16^6-17^5	111	23.61	2.80	24.63	3.99
17^6-18^5	110	24.26	4.17	23.99	5.15
18^6-19^5	93	24.59	3.78	25.36	3.90

portrayed in Figures 1 and 2. The continuous line connects the actual mean scores at each age level, whereas the broken line represents the smoothing of the data, and is probably the better indicator of trends. In general, the curve shows a decrease in illusion score with increase in age, a decrease which, from the results of the analysis of variance, is found to be statistically significant at the one per cent level. A scrutiny of the curve, however, reveals much more information concerning the growth phenomena than does the statistical test of significance. It will be observed that from ages six to 11 there is a consistent decrease in mean illusion score, followed by an increase which culminates at age 14, and which is followed by a slight decrease up to age 17. The curve from ages 17 to 19 shows a very slight increase. The decrease in ability to discriminate demonstrated about the age of puberty parallels the results obtained by Gilbert (22) in his study of a number of factors, such as muscle sensitivity, sensitivity to color differences, and the like. Gilbert

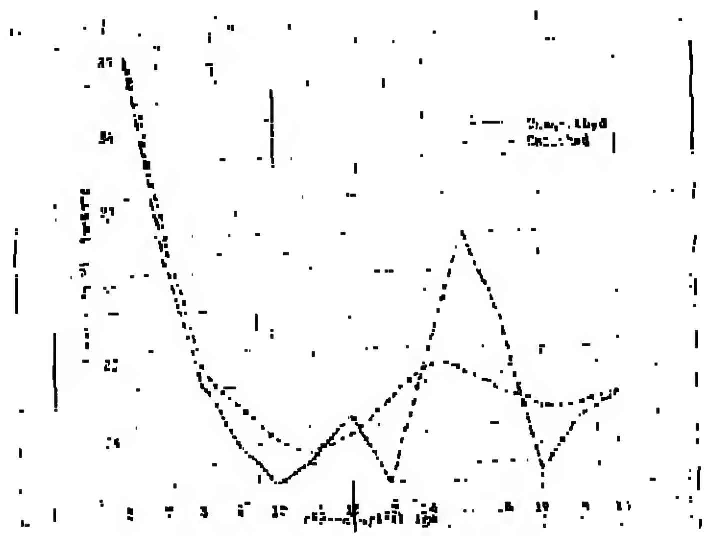

FIGURE 1
MEAN MÜLLER-LYER ILLUSIONS SCORES AT DIFFERENT AGE LEVELS

concluded from his data that the "children appear to have labored under some disadvantage in almost all tests at the period about 13." In the Müller-Lyer data of this experiment, the disadvantage appeared somewhat later.

It is important to note, at this point, that the generalized growth curve for the Müller-Lyer illusion, as described above, may be very misleading if account is not taken of sex differences after the age of puberty. The curve for boys and the curve for girls differ markedly during this period, as shown in Figures 13-16. The generalized curve of Figure 1, therefore, should be regarded merely as the midpoint between the means of the two sexes.

2. *Vertical-Horizontal Illusion*

The means and standard deviations of vertical-horizontal scores for ages six through 19 are given in Table 5 and are graphically portrayed in Figures 3 and 4. The smoothed curve indicates a consistent decrease in illusion from ages six to 17, with a rather marked increase from ages 17 to 19. The decrease from ages six to eight is almost imperceptible and is followed by a sharp decrease

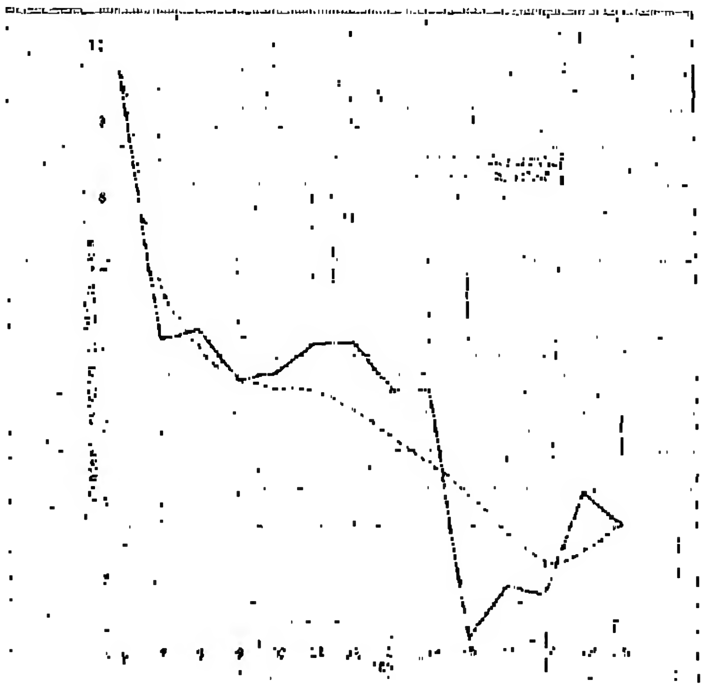

FIGURE 2
STANDARD DEVIATIONS FOR THE MÜLLER-LYER ILLUSION AT DIFFERENT AGE LEVELS

from ages eight to 11. The decrease becomes more gradual from ages 11 to 14, and is relatively great between ages 14 to 17. The relatively slow rate of decrease during the period from 11 to 14 is suggestive of the effects of puberty. Apparently the effects of puberty on this illusion are not so marked as they were for the Müller-Lyer illusion. The mean age differences are found, by the analysis of variance technique, to be statistically significant at the one per cent level.

B. MEAN VARIATIONS AROUND INDIVIDUAL MEANS AT DIFFERENT AGE LEVELS

1. *Müller-Lyer Illusion*

The means of mean variations around individual means are given

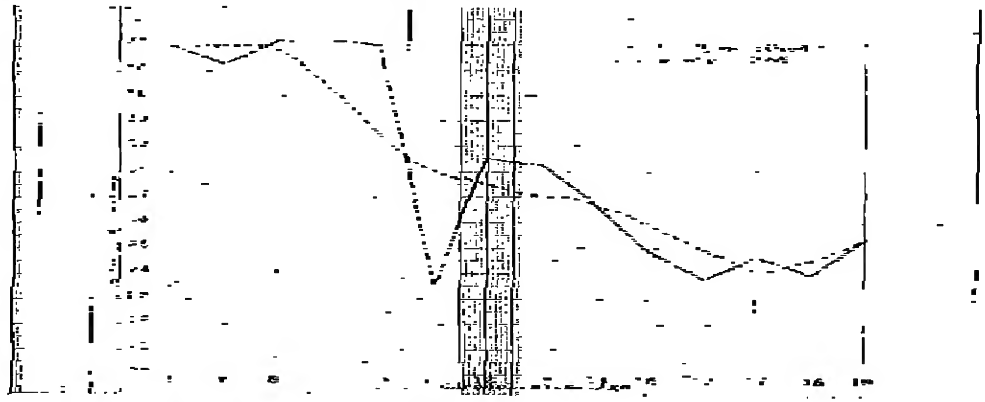

FIGURE 3
MEAN VERTICAL-HORIZONTAL ILLUSION SCORES AT DIFFERENT AGE LEVELS

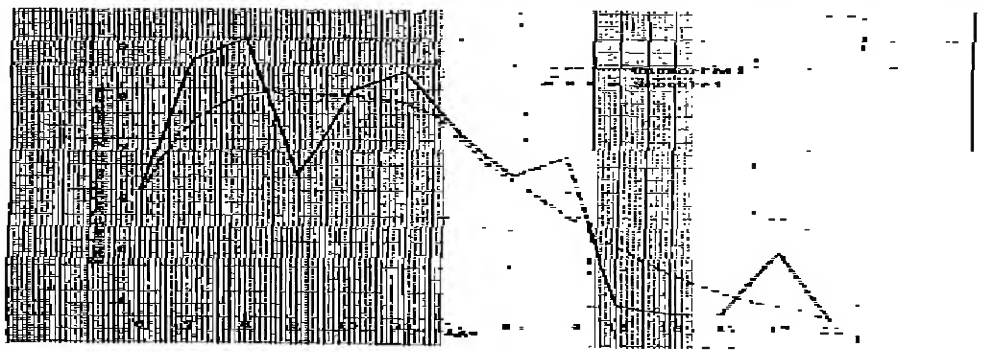

FIGURE 4
STANDARD DEVIATIONS FOR THE VERTICAL-HORIZONTAL ILLUSION AT DIFFERENT AGE LEVELS

TABLE 6

MEANS OF MEAN VARIATIONS AROUND INDIVIDUAL MEANS

Age group	M.-L. Illusion	V.-H. Illusion
5^{0}– 6^{6}	9.31	10.59
6^{6}– 7^{6}	9.63	12.36
7^{0}– 8^{6}	10.58	12.68
8^{0}– 9^{6}	10.35	13.28
9^{0}–10^{6}	10.67	11.28
10^{0}–11^{6}	12.98	13.94
11^{0}–12^{6}	9.88	10.52
12^{0}–13^{6}	8.79	10.45
14^{0}–15^{6}	8.56	9.91
15^{0}–16^{6}	9.02	8.49
16^{0}–17^{6}	8.10	7.17
17^{0}–18^{6}	7.41	6.78
18^{0}–19^{6}	8.37	9.12

in Table 6. It will be observed that the curve for mean variation at the different age levels, as shown in Figure 5, runs directly counter to the curve for mean illusion score up to the age of 14. From this age upward, the two curves run parallel. Variations increase

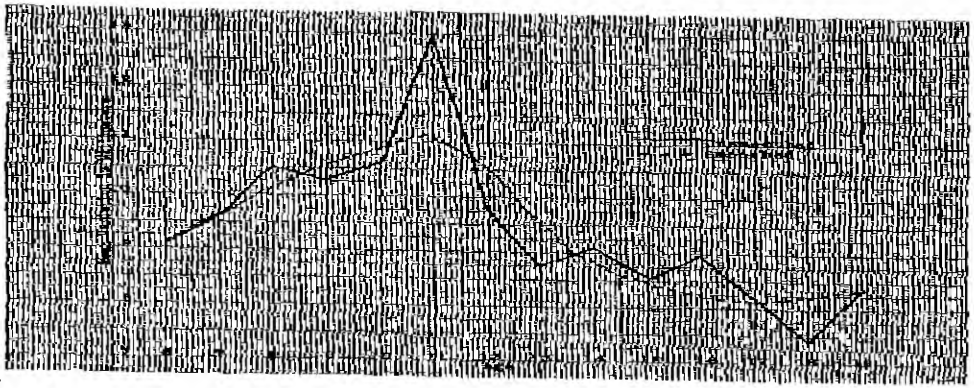

FIGURE 5

MEANS OF MEAN VARIATIONS AROUND INDIVIDUAL MEANS AT DIFFERENT AGE LEVELS (MÜLLER-LYER ILLUSION)

steadily from ages six to 11, reaching the highest point at age 11, which is the age at which the lowest mean illusion score is obtained. As variation decreases from ages 11 to 14, the corresponding mean illusion scores increase.

It is of interest at this point to compare the above results with those obtained in other studies of psychophysical functions. Gilbert (22), for example, in studying the muscle sense found that, on the whole, variation decreases with advance in age, and that when discriminative ability decreases between any two successive ages, the variation increases for the corresponding period. The Müller-Lyer data of this study show just the opposite trend, and are more consistent with the results of Gilbert's time-memory and suggestion tests. In the former test he discovered that where a special change in relative growth between two ages occurs for the better, the mean variation changes for the worse. In the latter test he found that for short periods in the development where ability increases, variation tends to increase also.

It is important at this point to recognize the effects of the marked sex differences in mean Müller-Lyer illusion upon the mean variations as described above. The fact that these differences are so great casts doubt upon the validity of generalizations drawn from the combined curve for the two sexes. These mean variations are given separately for each sex in Table 21 and are graphically shown in Figures 17 and 18. A description of these curves is given in the section on sex differences.

2. *Vertical-Horizontal Illusion*

The curve for mean variation obtained on the vertical-horizontal illusion, as shown in Figure 6, reveals a marked increase from ages six to eight, a decrease from ages eight to 17, and a marked increase again from ages 17 to 19. The increase in variation from ages six to eight runs parallel to mean illusion scores that decrease at less than the average rate of decrease from one age to the next. A similar trend was observed by Gilbert in his study of reaction time, and may be summarized as follows: Whenever the rate of increase in ability from one age to another is less than the average rate of increase, the variation for that period becomes higher and therefore worse. The curve for decrease in variation from ages eight to 17 may be characterized as positively accelerated and does not show

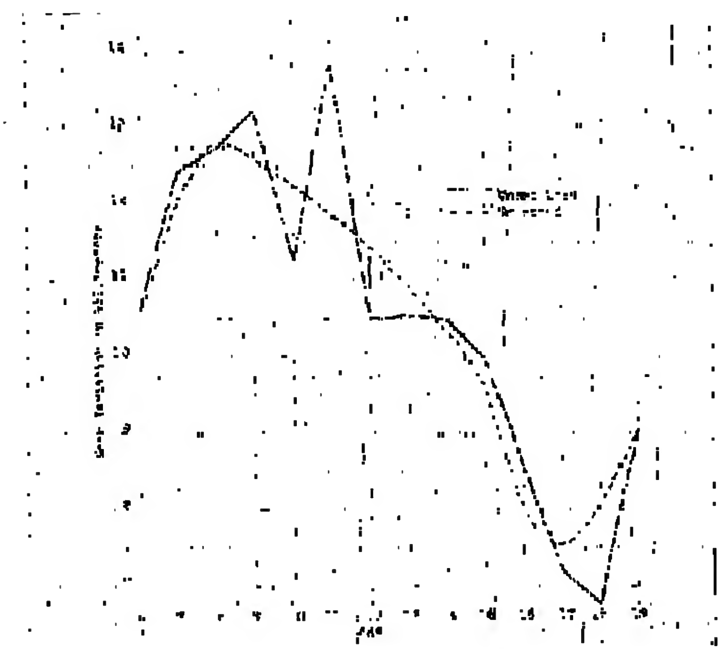

FIGURE 6
MEANS OF MEAN VARIATIONS AROUND INDIVIDUAL MEANS (VERTICAL-HORIZONTAL ILLUSION)

the disturbance at puberty which is characteristic of the curve for mean illusion score.

C. Trends in Reliability

Table 1 gives split-half reliability coefficients for the Müller-Lyer illusion corrected by the Spearman-Brown formula. It will be noted that the reliability coefficient for age six is considerably lower than that for age seven and the other ages up to 19. From ages seven to 19 there appears to be a very slight and statistically insignificant trend toward higher reliabilities with increasing age. These trends suggest that developmental changes are taking place, the most marked of which occur between ages six and seven. The low reliability at age six is probably indicative of what Luria (23) describes as a state of diffused nervous excitation, characteristic of

the young child, but which in the normal child has pretty well disappeared by the age of seven or eight years and has been replaced by the more constant integrated reaction which is characteristic of the adult. Goodenough's (23) study of reaction time led to findings in agreement with this. Goodenough holds that the fluctuations from one trial to another, so characteristic of the young child, cannot be attributed to distractions from external sources. They result rather from the child's inability to give an integrated response, in spite of the fact that his attention is centered on doing so.

The reliability coefficients for the vertical-horizontal illusion, as shown in Table 2, do not increase with advance in age, but are quite stable throughout the entire age range.

D. Distribution Curves for Illusion Scores at Different Age Levels

1. *Müller-Lyer Illusion*

It will be seen by reference to Table 7 that the distributions of Müller-Lyer illusion scores for age groups 7, 8, 9, and 10 are

TABLE 7
DATA ON NORMALITY OF DISTRIBUTIONS OF MÜLLER-LYER SCORES

Age group	g_1	SE_{g_1}	t	g_2	SE_{g_2}	t
5^{6}- 6^{5}	—.003	.5363	.005	.029	1.04	.278
6^{6}- 7^{5}	.983	.2346	4.18*	1.49	.4652	3.20*
7^{6}- 8^{5}	2.19	.2172	10.09*	3.91	.4314	9.07*
8^{6}- 9^{5}	.800	.2413	3.32*	2.45	.4784	5.13*
9^{6}-10^{5}	1.08	.2076	5.19*	1.97	.4113	4.79*
10^{6}-11^{5}	.050	.1936	.258	1.08	.3850	2.81†
11^{6}-12^{5}	— .090	.1876	— .480	— .09	.3735	— .241
12^{6}-13^{5}	.267	.2076	—1.28	.313	.4113	— .762
13^{6}-14^{5}	—.123	.2191	— .562	— .25	.4349	— .575
14^{6}-15^{5}	—.2603	.2337	—1.11	1.47	.4633	3.17*
15^{6}-16^{5}	.1069	.2368	.4514	.0391	.4694	.0832
16^{6}-17^{5}	.0924	.2305	.4008	1.08	.4571	2.36
17^{6}-18^{5}	.0704	.2526	.2787	— .1029	.5002	— .2057
18^{6}-19^{5}	.3633	.2540	1.43	.0369	.4577	.0806

*Significant at the one per cent level.
†Significant at the five per cent level.

positively skewed, and that age groups 7, 8, 9, 10, and 15 depart significantly from normality at the one per cent level in the direction of peakedness, while age group 11 is significantly peaked at the

five per cent level. The significance of these data for the analysis of growth will now be discussed.

It has usually been assumed that psychophysical functions are normally distributed in a random sample of the population, although noteworthy exceptions have been observed. Curves which appear to be normal may prove to be skewed when mathematical measures of goodness of fit are applied (8, 20, 26, 54, 82). Boring (8) in particular, has been especially insistent that a normal distribution should not be assumed a priori in psychophysical measurement until experimental evidence to that effect has been collected. So far as the writer knows, the present experiment is the only study in which a sufficiently large number of cases has been available for testing the hypothesis that susceptibility to the Müller-Lyer illusion is normally distributed.

It will be observed that the illusion scores are symmetrically distributed at each age from 11 through 19, whereas from ages seven through 10 they are positively skewed. In other words, there is a tendency at the lower age levels for illusion scores to pile up at the higher end of the scale. A comparison of Table 7 with Figure 1 suggests that skewness and change in susceptibility to illusion are related, that is, skewed distributions are obtained for these age groups where growth is most rapid. It seems reasonable to assume that in a function as complex as that presented by the Müller-Lyer illusion, the child may not be sufficiently "structuralized" to react adequately until the age of 11. The peakedness at age 11 and at age 15 may possibly be evidences of transition periods in growth.

It will be of interest at this point to note that skewness as a function of growth has been observed in certain biological data. Boring (8) calls attention to Raymond Pearl's statistical study[1] of the number of leaves in a whorl in ceratophyllum, in which the distributions for number of leaves were skewed one way at the proximal end of the main stem, and the other way at the distal end. The form of the distribution passed through symmetry somewhere between these two ends. In view of these facts, Pearl wrote:

> The phenomenon of skew variation stands forth in this case, free of doubtful interpretation through selection or any similar factor, clearly and definitely as a phenomenon of growth. In the face of facts of this kind it is difficult to understand how anyone can be so convinced of the "*Allgemeingültigkeit*" of the

[1]*Variation and Differentiation in Ceratophyllum, 1907.*

> normal or Gaussian law, as some biologists still are. . . . Skewness *in variation is a very real* biological phenomenon which may be changed and modified, not only in degree, but in direction, by various biological factors like growth, as for example, in the present case.

An alternative to the hypothesis discussed in the above paragraphs is, however, suggested by a scrutiny of Figure 9. This graph portrays the distribution of means of individual Müller-Lyer illusion scores based upon four and upon 20 trials for a group of 50 children, average age seven years and four months. It will be seen that the distribution of scores based upon 20 trials is more symmetrical than that based upon four trials, a fact which suggests that skewness may be a function of the unreliability of the measurements. To test this hypothesis, Fisher's g_1 and g_2 statistics were computed, together with their standard errors and t-values, as shown in Table 8. It will be observed that neither g_1 nor g_2 differ significantly from zero when the distributions are based upon 20 trials, and that the curve therefore meets the criteria for normality.

TABLE 8
DATA ON NORMALITY OF ILLUSION SCORES BASED ON 20 TRIALS (50 SUBJECTS, AVERAGE AGE SEVEN YEARS AND FOUR MONTHS)

Illusion	g_1	SE_{g_1}	t	g_2	SE_{g_2}	t
Müller-Lyer	—.0083	.3367	.0246	.0717	.6618	.1084
Vertical-Horizontal	.0700	.3367	.2079	—.9002	.6618	1.3602

2. *Vertical-Horizontal Illusion*

The vertical-horizontal illusion is symmetrically distributed at all age levels except 10, where it is positively skewed at the one per cent level of significance, and 15, where it is positively skewed at the five per cent level of significance as shown in Table 9. It will be noted that these are ages at which changes in mean illusion scores appear, as shown in Figure 2.

Values of t for g_2 are significant at the one per cent level for ages 7, 8, 10, 12, and 18, and at the five per cent level for age 14. In spite of this fact, however, the distributions appear to resemble normal curves more than they resemble any other common type of curve.

The skewness observed in the vertical-horizontal data may, like

TABLE 9

DATA ON NORMALITY OF DISTRIBUTION OF VERTICAL-HORIZONTAL ILLUSION SCORES

Age group	g_1	SE_{g_1}	t	g_2	SE_g	t
5^{0}- 6^{5}	— .3337	.5363	.622	1.28	1.04	1.23
6^{0}- 7^{5}	.286	.2346	1.22	1.73	.4652	3.72*
7^{0}- 8^{5}	.184	.2172	.848	2.73	.4314	6.33*
8^{0}- 9^{5}	— .346	.2413	—1.44	— .13	.4784	— .272
9^{0}-10^{5}	2.19	.2903	10.48*	2.20	.4143	5.31*
10^{0}-11^{5}	.255	.1936	1.31	.60	.3850	1.56
11^{0}-12^{5}	.276	.1876	1.47	1.51	.3735	4.04*
12^{0}-13^{5}	.230	.2076	1.11	.45	.4113	1.09
13^{0}-14^{5}	.641	.2191	2.93	1.15	.4349	2.64†
14^{0}-15^{5}	.622	.2336	2.661†	.789	.5631	1.704
15^{0}-16^{5}	— .174	.2368	— .7331	.055	.4694	.1173
16^{0}-17^{5}	.504	.2305	2.1876	.986	.4571	2.1559
17^{0}-18^{5}	.290	.2526	1.1480	—1.29	.5002	—2.5789
18^{0}-19^{5}	— .023	.2413	— .0957	2.34	.4784	4.8912*

*Significant at one per cent level.
†Significant at five per cent level.

that for the Müller-Lyer, be attributed to the unreliability of mean individual scores based upon four trials. Reference to Table 8 indicates that neither g_1 nor g_2 differ significantly from normality when scores are based upon 20 trials.

E. LONGITUDINAL STUDY

1. *Selection of Subjects*

As a check against the results of the cross-sectional study, individuals from age groups 6, 7, 8, and 9 were tested a second time, after a period of three years had elapsed. All children of these age groups who were in the same school three years after the original testing were included in the follow-up groups. The number of subjects in each group is shown in Table 4.

2. *Correlations between First and Second Testings*

The correlations between initial scores and final scores on the two illusions are given in Table 10. It will be noted that correlations for the Müller-Lyer illusion, with the exception of age group six, show a consistent increase with increasing age, although the only statistically significant correlation is at age eight. If we exclude the high correlation at age six, which may be attributed to the small number of cases, we find a consistent trend toward higher correla-

TABLE 10

CORRELATION BETWEEN INITIAL SCORES AND FINAL SCORES (THREE YEARS LATER) ON MÜLLER-LYER AND VERTICAL-HORIZONTAL ILLUSION

Age: First testing	Age: Second testing	No.	Müller-Lyer r	Müller-Lyer σr	Vertical-Horizontal r	Vertical-Horizontal σr
5^0-6^5	8^0- 9^5	18	.5272	.3224	.1675	.4334
6^0-7^5	9^0-10^5	70	.0351	.1194	*.4322	.0972
7^0-8^5	10^0-11^5	83	.2357	.1036	*.3697	.0947
8^0-9^5	11^0-12^5	63	*.5945	.0814	*.4209	.1037

*Statistically significant correlation.

tions with advance in age, a trend which suggests a growth factor. Apparently, reaction to the Müller-Lyer illusion becomes a more stable function with increasing age. This may be regarded as an additional line of evidence that the child becomes more "structuralized" with advance in age.

The correlations for the vertical-horizontal illusion show a marked increase from ages six to seven, after which they remain remarkably stable and are statistically significant. This sudden rise in the magnitude of the correlation between ages six and seven is suggestive of a change in growth.

A comparison of the correlations for the two illusions suggests that reactions to the vertical-horizontal illusion become stabilized at an earlier age than do reactions to the Müller-Lyer. This is consistent with other data, such as the previously reported reliability coefficients which reveal a similar trend.

3. *Mean Differences between First and Second Testing*

a. Müller-Lyer Illusion. The mean differences in Müller-Lyer illusion scores between the first and second testing, and the significance of these differences are given in Table 11. It will be observed that these differences are statistically significant for all except the youngest age group. It will also be seen that the direction of these differences is opposite to that obtained in the cross-sectional study, a result which was quite unexpected. A discussion of this phenomenon is given in the section on growth increments and decrements.

b. Vertical-Horizontal Illusion. The mean differences in vertical-horizontal illusion scores between the first and second testing, and the significance of these differences are shown in Table 12. It will be observed that the direction of the differences is the same as that

TABLE 11
MEAN DIFFERENCES BETWEEN INITIAL AND FINAL SCORES ON THE MÜLLER-LYER ILLUSION

Test 1	Test 2	N	Mean$_1$	σ	σ_{m1}	Mean$_2$	σ	σ_{m2}	*Diff.*	$\sigma_{diff.m_1-m_2}$	t
5^6-6^5	8^6-9^5	18	27.40	6.94	2.1419	31.90	5.98	1.8456	+4.50	1.9555	2.301†
6^6-7^5	9^6-10^5	70	28.85	6.03	.720	30.48	4.79	.572	+3.63	.9046	4.012*
7^6-8^5	10^6-11^5	83	24.53	6.72	.7475	30.41	5.91	.6490	+5.88	.866	6.097*
8^6-9^5	11^6-12^5	63	24.40	5.38	.6776	29.23	5.21	.6569	+4.83	.601	8.036*

*Statistically significant at one per cent level.
†Statistically significant at five per cent level.

TABLE 12
MEAN DIFFERENCES BETWEEN INITIAL AND FINAL SCORES ON THE VERTICAL-HORIZONTAL ILLUSION

Test 1	Test 2	N	Mean$_1$	σ	σ_{m1}	Mean$_2$	σ	σ_{m2}	*Diff.*	$\sigma_{diff.m_1-m_2}$	t
5^6-6^5	8^6-9^5	18	35.60	10.53	3.250	28.90	7.53	2.316	—6.70	3.661	1.830
6^6-7^5	9^6-10^5	70	32.43	8.01	.957	29.99	7.19	.859	—2.44	.9710	2.512†
7^6-8^5	10^6-11^5	83	32.43	9.56	1.049	30.01	7.304	.802	—2.42	1.063	2.276*
8^6-9^5	11^6-12^5	63	33.33	6.11	.770	30.67	6.73	.848	—2.66	.8726	3.048*

†Statistically significant at five per cent level.
*Statistically significant at one per cent level.

obtained in the cross-sectional study. Differences between ages six and nine are not statistically significant, between ages seven and 10 and between ages eight and 11 they are significant at the five per cent level, and between ages nine and 12 they are significant at the one per cent level.

4. *Growth Increments and Decrements*

a. Müller-Lyer Illusion. Mean individual growth increments for the Müller-Lyer illusion over the three-year period, as shown in Table 13 are positive for all age groups, and about as great for one age group as for another. This is in line with the direction

TABLE 13
GROWTH INCREMENTS FOR MÜLLER-LYER ILLUSION (AFTER THREE-YEAR INTERVAL)

Age	Mean increments	σ	σ_m	Significance 5% level	1% level
5^0-6^5 and 8^0-9^5	4.89	7.81	1.84	3.88	5.32
6^0-7^5 and 9^0-10^5	4.19	7.38	.88	1.75*	2.37*
7^0-8^5 and 10^0-11^5	6.19	7.89	.87	1.71*	2.28*
8^0-9^5 and 11^0-12^5	4.80	7.12	.90	1.80*	2.39*

*Statistically significant.

of the mean group differences for the follow-up group but is inconsistent with the results of the cross-sectional study. The explanation of this discrepancy is not clear. The equality of the growth increments for the different age groups suggests that a constant factor is operating in every case. The first possibility that suggested itself was that an experimental error had been inadvertently introduced into the situation at the time of the second testing. However, the writer is at a loss to discover any such sources of error, since the conditions of the experiment were identical for the two testings. Both tests were given by the same examiner in the same room, and with the same apparatus. The only difference between the children in the follow-up group and the children of the same age in the cross-sectional study was that the former had been exposed to the illusion three years earlier. A scrutiny of Figure 7, which shows the mean scores of all four groups on the two tests, shows a consistent trend downward from the younger to the older groups. Hence, although all groups have increased in illusion score over a three-year period,

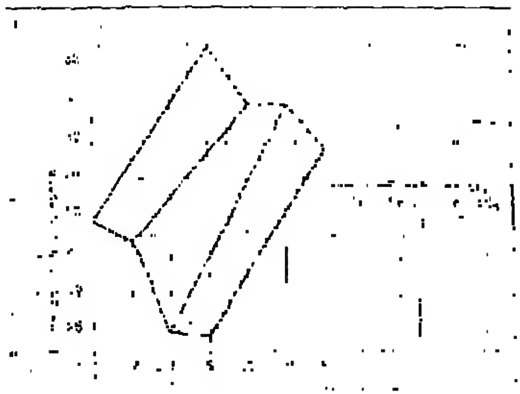

FIGURE 7

COMPARISON OF MÜLLER-LYER ILLUSION SCORES FOR FIRST AND SECOND TESTING (THREE-YEAR INTERVAL)

the relative position of the means remains the same, and is consistent with the results of the cross-sectional study.

b. Vertical-Horizontal Illusion. Mean individual growth decrements for the vertical-horizontal illusion over a three-year period, and the significance of these differences are shown in Table 14. It

TABLE 14

GROWTH DECREMENTS FOR VERTICAL-HORIZONTAL ILLUSION (AFTER THREE-YEAR INTERVAL)

Age	Mean decrements	σ	σ_m	Significance 5% level	1% level
5^0-6^5 and 8^0- 9^5	—6.55	10.86	2.56	2.54	7.40
6^0-7^5 and 9^0-10^5	—2.41	9.47	1.13	2.25*	2.97
7^0-8^5 and 10^0-11^5	—2.43	9.63	1.06	2.11*	2.79
8^0-9^5 and 11^0-12^5	—2.65	7.11	.90	1.80*	2.39*

*Statistically significant.

will be observed that the direction of the differences is the same as that obtained in the cross-sectional study, although, in general, the differences are not so great. Differences between ages six and nine are not statistically significant, between ages seven and 10 and between ages eight and 11 they are significant at the five per cent level, and between ages nine and 12 they are significant at the

one per cent level. The mean differences between illusion scores on the first and second testing are graphically shown in Figure 8.

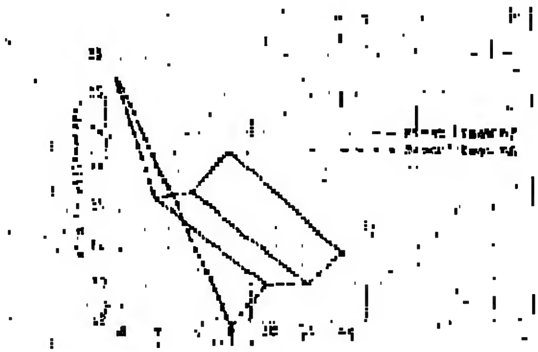

FIGURE 8
COMPARISON OF VERTICAL-HORIZONTAL ILLUSION SCORES FOR FIRST AND SECOND TESTING (THREE-YEAR INTERVAL)

It will be observed that, with the exception of the youngest age group, the relative position of the group means remains the same for both tests.

5. *Effect of Exposure to Illusion Over a Short Period of Time*

a. Müller-Lyer Illusion. The discrepancy between the results of the cross-sectional and of the longitudinal study suggests that a few exposures to the illusion may increase the illusory effect, although, as previous research has demonstrated, many exposures tend to reduce it. An additional experiment was, therefore, set up to determine what effect exposure to four trials would have on illusion scores secured sometime later.

A group of 42 first- and second-grade children, average age seven years and four months, was given a series of four trials, and was retested under identical conditions from five to eight weeks later. The mean scores for the first and second series of trials, the standard deviations, and the significance of differences are given in Table 15. It will be observed that the mean Müller-Lyer illusion increases 3.95 millimeters on the second test, an increase which is significant at the five per cent but not at the one per cent level of significance. This trend is in the same direction as the trend obtained in the three-

TABLE 15

MEAN DIFFERENCES BETWEEN ILLUSION SCORES SECURED OVER AN INTERVAL OF FROM FIVE TO EIGHT WEEKS

Illusion	$Mean_1$	σ	σ_{m1}	$Mean_2$	σ	σ_{m2}	Diff.	$diff._{m_1-m_2}$	t
Müller-Lyer	30.68	8.942	1.4139	34.63	6.904	1.1363	1.75	.8979	1.95
Vertical-Horizontal	33.74	7.073	1.1184	35.29	7.186	1.0917	3.95	1.9306	2.04†

†Statistically significant at five per cent level.

year follow-up study, although it is not of such pronounced statistical significance. The mean of the individual increments together with the standard errors and t-values, are given in Table 16. It will be

TABLE 16
MEAN INCREMENTS IN ILLUSION SCORE (AFTER FIVE TO EIGHT WEEKS' INTERVAL)

Illusion	Mean *increment*	σ	σ_m	t
Müller-Lyer	3.76	1.4272	.2203	17.5*
Vertical-Horizontal	1.78	.9065	.1553	11.4*

*Significant at the one per cent level. $N = 42$

seen that this mean increment is statistically significant at the one per cent level.

b. Vertical-Horizontal Illusion. The vertical-horizontal illusion was then presented to the same group of children that had been exposed to the Müller-Lyer illusion, and under the same conditions. The data for the first and second series of trials with this illusion are shown in Table 15. Although there is a positive increment in mean illusion score, it is not statistically significant. The mean of the individual increments as well as the standard errors and t-values are given in Table 16. The mean of the individual increments is statistically significant at the one per cent level.

F. RELATIONSHIP BETWEEN MÜLLER-LYER AND VERTICAL-HORIZONTAL ILLUSION

Product-moment coefficients of correlation between Müller-Lyer and vertical-horizontal illusions, corrected for attenuation, are given in Table 17. It will be noted that in general there is a low positive correlation between the two illusions, a result which corroborates an earlier study by Tinker (82). The corrected coefficients are statistically significant at all ages except 10, 12, 14, 15, and 16, and show a slight tendency to be higher at the lower age levels than at the upper age levels.

G. EXPERIMENTAL TEST OF RELIABILITY OF MEASUREMENTS

At this stage of the study it was thought desirable to give further consideration to possible sources of unreliability in the original measurements. It will be recalled that the mean illusion score for

TABLE 17

COEFFICIENTS OF CORRELATION FOR VERTICAL-HORIZONTAL AND MÜLLER-LYER ILLUSIONS AT EACH AGE LEVEL, CORRECTED FOR ATTENUATION

Age group	r	σr	r_{∞}
5- 6	.1736	.086	.3295*
6- 7	.2238*	.092	.2753*
7- 8	.2731*	.022	.3309*
8- 9	.3165*	.0001	.3716*
9-10	.0906	.078	.1066
10-11	.2303*	.037	.2579*
11-12	.1393	.061	.2579*
12-13	.2162*	.045	.2479*
13-14	.1158	.078	.1374
14-15	.1061	.085	.1246
15-16	—.0867	.097	—.0927
16-17	.1821	.063	.1976*
17-18	.2161*	.051	.2431*
18-19	.2268*	.050	.2461*

*Statistically significant.

each individual was based on four trials, and that the reliability coefficients were, with the exception of that obtained for the youngest age group, sufficiently high for making group comparisons. At most ages, however, they were not sufficiently high to insure the reliability of individual means. Therefore, means secured from four trials were correlated with means secured on 20 trials, for a group of 50 children from the first and second grade, average age seven years and four months. Split-half reliability coefficients based on 20 trials, and corrected by the Spearman-Brown formula, are given in Table 18. It will be seen that they are sufficiently high

TABLE 18

SPLIT-HALF RELIABILITY COEFFICIENTS BASED ON TWENTY TRIALS, CORRECTED BY SPEARMAN-BROWN FORMULA (AVERAGE AGE OF FIFTY SUBJECTS, 7^{4})

Müller-Lyer		Vertical-Horizontal	
r	r_{∞}	r	r_{∞}
.9128	.9546	.9456	.9721

to permit individual prediction. Müller-Lyer means based on four trials correlated .4322 with means secured on 20 trials, and vertical-horizontal means based on four trials correlated .9156 with means based on 20 trials. Since the four trials are a part of the 20 trials, however, this correlation is spuriously high. A part score based

on four trials in a total of 20 trials would theoretically have one-fifth of the determining factors in producing the total score. Since all the determining factors of the part are included within the determining factors of the whole, r^2 would equal one-fifth or .20, and r would equal .477, even though the part were actually unrelated to the remainder of the whole. In the case of the Müller-Lyer illusion, the removal of this spurious factor would reduce the correlation to zero. There is reason, therefore, for doubting the reliability of the Müller-Lyer means based upon four trials. The reliability of the vertical-horizontal means based upon four trials is much higher than the Müller-Lyer means, although this correlation is considerably lowered when the effect of the spurious factor is removed. A comparison of mean scores based on four trials and mean scores based on 20 trials is given in Figures 9 and 10. A similar comparison for mean variations is given in Figures 11 and 12.

H. Sex Differences

The mean illusion scores at different ages are given separately

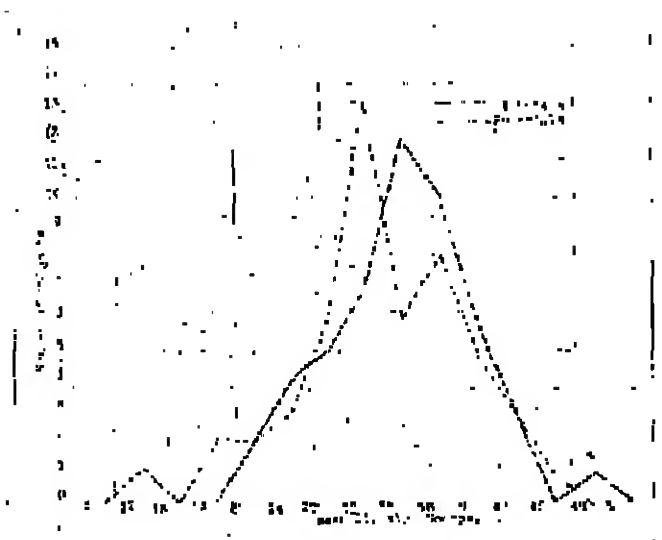

FIGURE 9

Comparison of Means of Individual Scores Based Upon Four and Upon Twenty Trials (Müller-Lyer Illusion)

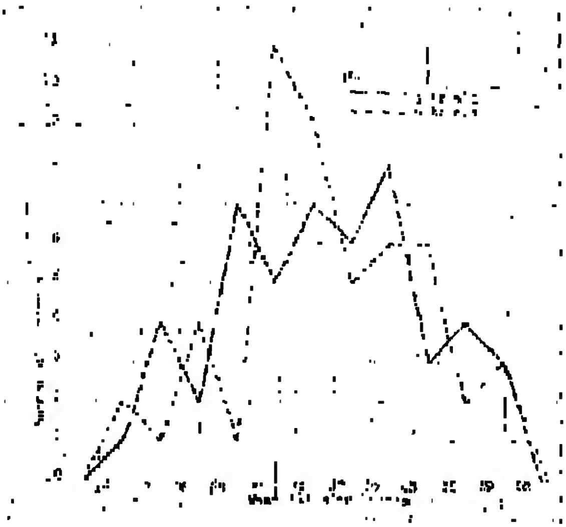

FIGURE 10

Comparison of Means of Individual Scores Based Upon Four and Upon Twenty Trials (*Vertical-Horizontal Illusion*)

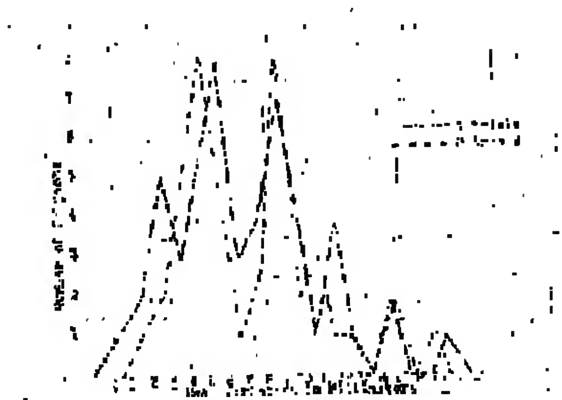

FIGURE 11

Comparison of Individual Mean Variations Based Upon Four and Upon Twenty Trials (Müller-Lyer Illusion)

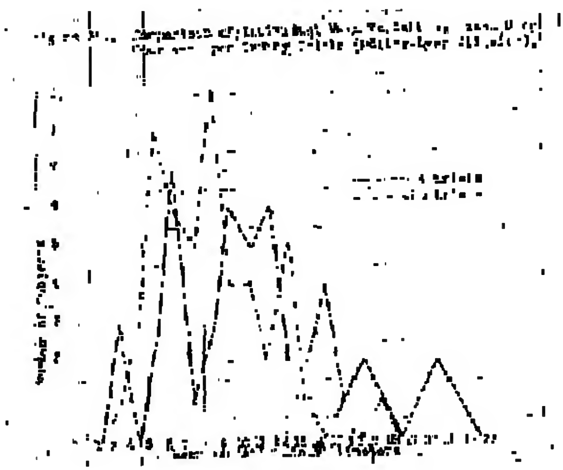

FIGURE 12

Comparison of Individual Mean Variations Based Upon Four and Upon Twenty Trials (Vertical-Horizontal Illusion)

TABLE 19

Mean Illusion Scores for Boys and for Girls at Different Age Levels

	Müller-Lyer Illusion				Vertical-Horizontal	
	Boys		Girls		Boys	Girls
Age group	*N*	Mean	*N*	Mean	Mean	Mean
5^0- 6^6	43	29.62	67	28.67	34.30	33.90
6^0- 7^6	54	28.56	52	26.48	32.60	32.06
7^0- 8^6	78	24.42	46	25.57	33.37	32.91
8^0- 9^6	45	25.75	55	22.35	34.36	32.06
9^0-10^6	71	24.78	66	22.07	32.99	31.59
10^0-11^6	88	24.11	73	22.84	32.18	29.22
11^0-12^6	77	24.77	90	23.75	29.36	27.81
12^0-13^6	62	24.02	77	22.83	29.25	27.50
13^0-14^6	48	23.84	74	25.53	28.04	26.24
14^0-15^6	12	22.85	96	27.20	30.02	24.33
15^0-16^6	38	20.44	66	28.26	24.60	23.66
16^0-17^6	68	21.66	43	26.70	25.34	23.58
17^0-18^6	50	21.27	60	26.76	24.81	22.40
18^0-19^6	76	25.37	17	21.75	25.55	24.56

for each sex in Table 19 and are graphically portrayed in Figures 13-16. Caution should be observed in interpreting these graphs at the upper age levels, since the number of subjects of each sex is

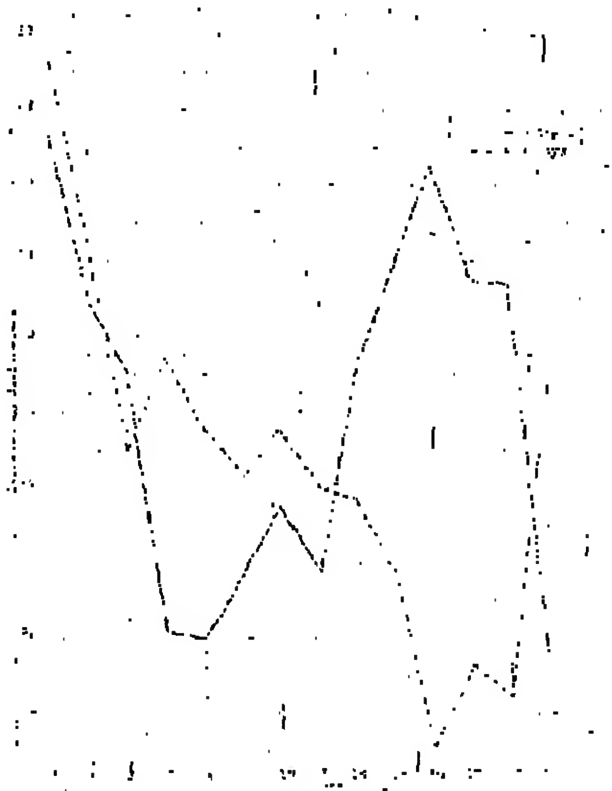

FIGURE 13
SEX DIFFERENCES IN MÜLLER-LYER ILLUSION

not constant. It should be noted in particular that the number of boys in age group 15 is only 12, and that the number of girls in age group 19 is only 17. These are considerably smaller numbers than are found in the other age groups.

1. *Müller-Lyer Illusion*

Definite sex differences appear at age nine, although the direction of the two curves is in general the same until age 11, as shown in Figure 13. From ages 11 to 17, the illusion for boys steadily

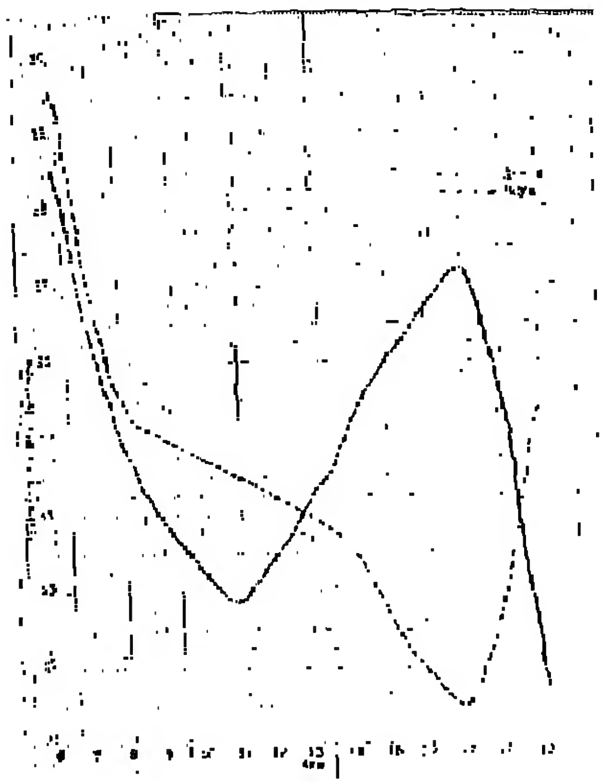

FIGURE 14
SEX DIFFERENCES IN VERTICAL-HORIZONTAL ILLUSION (SMOOTHED)

decreases, whereas that for girls steadily increases, and the two curves cross at about age 13 and again at about 18.

These tremendous sex differences manifested during adolescence require some explanation. It would seem more reasonable to attribute them to differences in attitude toward the experiment than to differences in perception, since differences in attitude were apparent during the testing. A number of boys, for example, stated that they knew the figure was an illusion, and volunteered explanations of the cause of the illusion. Several of them said it was a "trick" and wanted to remain to see what the other boys would do with it.

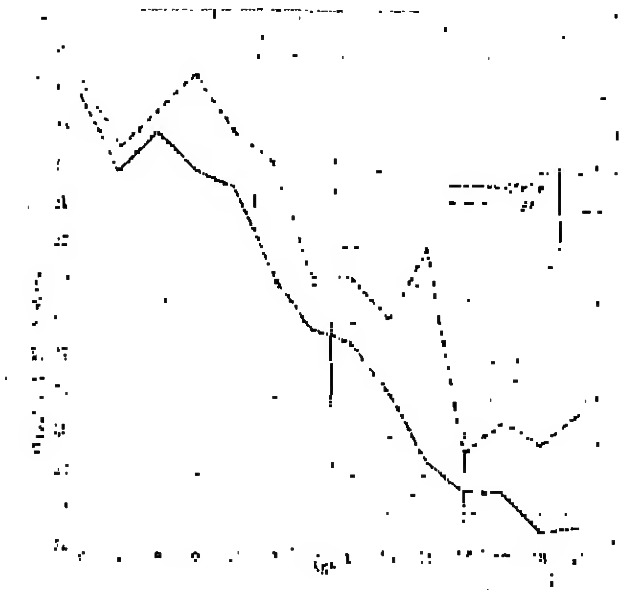

FIGURE 15
SEX DIFFERENCES IN VERTICAL-HORIZONTAL ILLUSION

None of the girls, however, made any comments about the situation, and although they took the experiment seriously, they showed very little interest in knowing their scores, and did not show the same eagerness and competitive spirit in taking the test as did the boys.

To test the hypothesis that the enormous sex differences observed are due, to some degree at least, to differences in attitude, an experiment was devised in which an attempt was made to arouse in the girls a critical attitude. The Müller-Lyer illusion was presented to 40 high school girls, average age 17 years and one month, and the following instructions were given:

> I want to see how accurately you can judge the equality of these two lines (point to the two lines) when you are distracted by the arrowheads (point) and the featherheads (point) that go out from them. These lines you see (point to featherheads and arrowheads) create an illusion so that the two lines look equal when they really are not. Little children are usually quite easily fooled by it. Older people, particularly the more

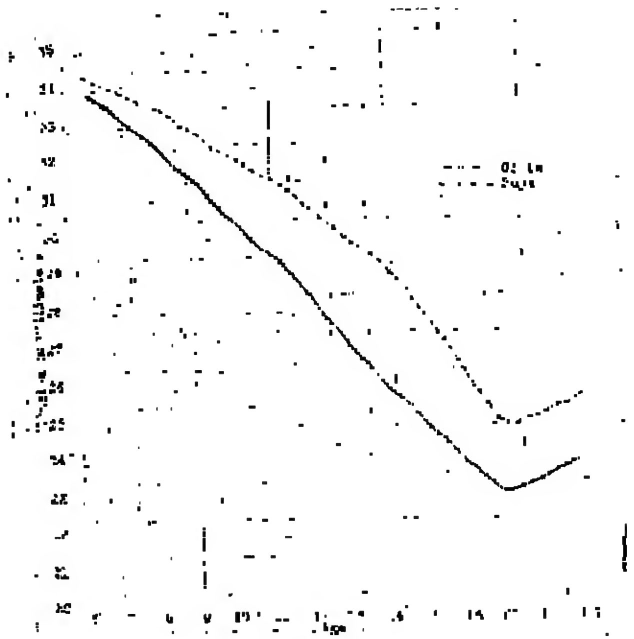

FIGURE 16

SEX DIFFERENCES IN VERTICAL-HORIZONTAL ILLUSION (SMOOTHED)

intelligent ones, of course, are not. Now I am going to move this line (point to the variable line) and I want you to tell me when it looks the same length as this (point to standard line) line. Be careful to concentrate so that you will not be deceived by the arrowheads and featherheads. I want to see which of you will get the best score!

The mean Müller-Lyer illusion score, obtained under the above conditions is 26.86. A comparison of this result with the previously secured result of 26.7, as given in Table 5, reveals no perceptible differences between the two means. It appears, therefore, that the directions given were inadequate for stimulating a more critical attitude, or else that the sex differences cannot be attributed to differences in attitude. The writer feels that it is extremely difficult to develop in girls a critical attitude toward a test situation such as this, since even with the added stimulation, the girls did not manifest the same concern over their scores as did the boys.

TABLE 20
SEX DIFFERENCES IN MEAN VARIATIONS AROUND INDIVIDUAL MEANS
(MÜLLER-LYER ILLUSION)

Age	Mean variation Boys	Girls
5^0- 6^5	9.56	9.05
6^0- 7^5	9.66	9.60
7^0- 8^5	11.01	10.16
8^0- 9^5	10.41	10.32
9^0-10^5	10.12	11.77
10^0-11^5	12.96	12.99
11^0-12^5	10.01	8.70
12^0-13^5	8.80	8.79
13^0-14^5	9.34	8.96
14^0-15^5	8.13	8.76
15^0-16^5	8.68	9.24
16^0-17^5	8.11	8.10
17^0-18^5	8.12	6.95
18^0-19^5	8.63	8.34

Sex differences in mean variation are given in Table 20, and are graphically shown in Figures 17 and 18. The curve for girls rises steadily from ages six to 11, and reaches its peak at age 11. From ages 11 to 17, the curve descends consistently, after which there is a very slight increase up to age 19. The highest mean variation occurs at age 11, which is the age at which the lowest mean illusion score is obtained. The lowest mean variation, on the other hand, occurs at age 17, which is the age at which the peak of the adolescent increase in mean illusion score is obtained.

The curve of mean variation for boys rises steadily from ages six to 11, and drops sharply from ages 11 to 14. From ages 14 to 17 there is a very slight drop, followed by a slight rise from ages 17 to 19. The lowest mean variation occurs at age 17, which is the age at which the lowest mean illusion is obtained.

2. *Vertical-Horizontal Illusion*

Sex differences in vertical-horizontal illusion are manifested at all age levels, as shown in Figures 15 and 16. The boys at all ages have a somewhat higher illusion score than do the girls, but the general trend of the two curves is the same. A possible explanation of the fact that boys do not have an advantage over girls during adolescence is that this figure does not present a "trick" as does

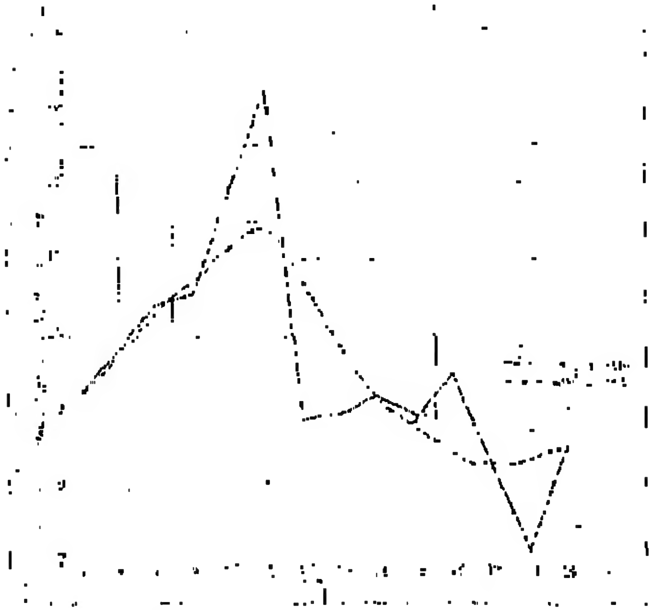

FIGURE 17
MEAN VARIATIONS AROUND INDIVIDUAL MEANS FOR MÜLLER-LYER ILLUSION (GIRLS)

the Müller-Lyer illusion, and consequently the attitude of the boys in taking the test is no more critical than is that of the girls.

I. INTELLIGENCE AND ILLUSION

The analysis of covariance yielded *F*-values indicating that the regression of personal constants on Müller-Lyer illusion is significant neither "within" nor "among" age groups. For the vertical-horizontal illusion the "among" regression is significant, but the "within" regression is significant at the one per cent level, and indicates a negative correlation between illusion scores and personal constants.

J. DISCUSSION

1. *Theories of Growth*

In accordance with the Gestalt explanation of the Müller-Lyer illusion, one would expect to find a progressive decrease in the amount

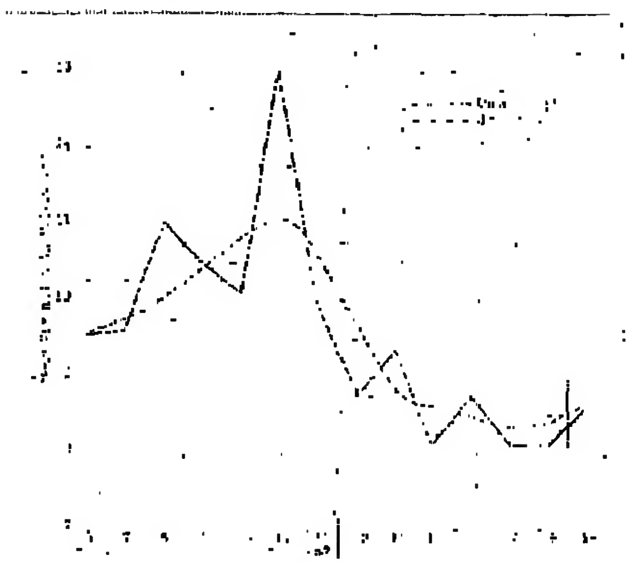

FIGURE 18
MEAN VARIATIONS AROUND INDIVIDUAL MEANS FOR MÜLLER-LYER ILLUSION (BOYS)

of illusion with advance in age until the maturity of the function is reached. It will be seen that the Müller-Lyer data of this experiment corroborate the Gestalt theory, inasmuch as the mean illusion decreases with increasing age up to 11 years. It seems probable that 11 years is the age at which maturity in this function is reached.

The Müller-Lyer data also yield evidence for Luria's (23) contention that "the 'simple reaction' in young children differs from the reaction of the adult in having another structure, and is characterized by a marked specificity of the diffused excitation, a weakness of those higher regulating mechanisms which are undoubtedly a basic phenomenon in the neurodynamics of the adult." The reliability coefficient, for example, was considerably lower for the youngest age group than for the succeeding age groups. To a lesser extent, this was also true for the vertical-horizontal illusion.

Further corroboration of Luria's thesis is found in a comparison by ages of the correlations between the two testings at three-year intervals. With increasing age the correlation coefficients went up,

a fact which indicates that the function is more stable at the higher age levels.

The phenomenon of skewness at the lower age levels has already received some consideration. It appears reasonable to suppose that the young child cannot react properly to such a complicated situation as that presented by the Müller-Lyer illusion because of a weakness of the "higher regulating mechanisms" as suggested by Luria.

A scrutiny of the coefficients of correlation between Müller-Lyer and vertical-horizontal illusions at the different age levels, suggests a somewhat closer mental organization at the lower age levels. The age differences, however, are so slight as to be almost negligible.

2. *Research Methods*

A number of methodological considerations are suggested by the present research. Considerable emphasis has recently been placed upon the use of a longitudinal approach in growth studies in preference to the cross-sectional approach that has been so commonly used. Shuttleworth, in particular, has urged not only the abandonment of cross-sectional data, but also the abandonment of statistical techniques developed for the analysis of cross-sectional data. This position, so far as anatomical and physiological data are concerned, appears to be well taken.

It has long been known, however, that many psychological processes change as a result of previous exposure to a stimulus, and hence, in testing such processes over a period of time, one gets not merely a measure of growth but also a measure of the effect of previous experience. The inconsistency of the results of the cross-sectional and of the longitudinal approach in the present study, as well as the results of the five- to eight-week follow-up study, are suggestive of this effect. In the light of such problems, it would appear unwise to follow Shuttleworth's suggestions indiscriminately when dealing with psychological processes, unless the effect of repeated measurements over a period of time has been definitely ascertained.

A few words concerning the use of statistical tests of significance in the analysis of growth data will here be in order. The method used in this study, namely, the analysis of variance technique, with its accompanying test of significance, is being increasingly used in psychological studies. It may be of interest, therefore, to point out

some distinct limitations of this technique for the analysis of cross-sectional growth data in general.

One of the basic assumptions underlying the analysis of variance technique is that the variances of the groups are homogeneous. In the present study this assumption was proved valid by means of Nayer's test (57), and the analysis of variance technique could therefore be used. There is reason for supposing, however, that for growth data in general, the variances will not be homogeneous, since changes in variance are usually an essential feature of the growth process. Jones (35), for example, in her study of the ability of preschool children to equate length of lines on the Galton bar, found that standard deviations decreased with age. Similarly, in other types of data, a decrease in the variability of individual scores is frequently accepted as a criterion of growth. In the light of such facts, it would appear unwise to use in the investigation of growth any technique of analysis which assumes that the variances among different age groups are homogeneous. If homogeneity should be demonstrated, however, as in the present research, such a technique might well be used. Nevertheless, most growth data will probably not yield homogeneous variances.

3. *Relation to Other Perceptual Phenomena*

It is reasonable to suppose that optical illusions are related to the phenomena of perceptual constancy, since in both cases the physical stimulus and the corresponding psychological reaction are at variance. It will therefore be of interest to compare the results of the present study with results of studies of perceptual constancy, as reviewed by Locke (52).

According to Locke, there is considerable evidence that perception and intelligence function within the organism as compensatory adjustive mechanisms. Consequently, as the organism grows more complex there is an increase in intelligence and a correlative decrease in the primitive perceptual mechanism. This is illustrated by experiments indicating that perceptual constancy decreases as one ascends the phylogenetic scale. Fish, chicks, and chimpanzees, for example, are superior to man, chicks are superior to chimpanzees and children are superior to adults. In general, it seems reasonable to suppose that geometrical-optical illusions follow a similar trend, since these illusions have been demonstrated in animals, and since children have these illusions to a greater degree than do adults. Further corrobora-

tion of such an hypothesis is given by the fact that "within" age groups, vertical-horizontal illusion scores correlated negatively with intelligence ratings.

A relationship between the common illusions and eidetic imagery has been suggested by Jaensch (33), and hence it will be of interest to compare age data on these two phenomena. Klüver (42) summarizes the results of eidetic imagery studies as follows:

> While there seems to be agreement that *EI* disappear, as a rule, during or after puberty, the period of the "acme" of eidetic imagery, the age at which *EI* are most frequently found, and found in the highest "degree," is not yet agreed upon. Most often the age of 12 years has been suggested; but, after Roessler's investigation, six years or below promises to be most correct.

If Jaensch's contention is true, one would expect to find, in the light of the above summary, that the greatest changes in susceptibility to illusion take place before the age of six. From our knowledge of perceptual development in general, this would appear to be a reasonable assumption.

IV. SUMMARY AND CONCLUSIONS

The two objectives of the present research were (*a*) to determine if susceptibility to certain geometrical-optical illusions varies significantly and predictably with age, and (*b*) to make an analysis of growth trends in the perception of such illusions. For this purpose, quantitative measures of Müller-Lyer and vertical-horizontal illusion were secured for 1,693 subjects, ranging in age from six to 19 years; and, of this group, 234 subjects at four age levels were selected for a follow-up study of three years. The principal results of this research are as follows:

1. Mean Müller-Lyer and vertical-horizontal illusion scores vary significantly "among" age groups, as determined by the analysis of variance and its accompanying test of significance.
2. The regression of personal constants on Müller-Lyer illusion scores is significant neither "within" nor "among" age groups.
3. The regression of personal constants on vertical-horizontal illusion scores is not significant "among" age groups. "Within" age groups it is significant and is negative in direction.
4. Mean Müller-Lyer illusion scores decrease consistently from ages six to 11, increase somewhat from ages 11 to 14, decrease slightly from ages 14 to 17, and increase very slightly from ages 17 to 19. Beyond the age of 11, however, the generalized growth curve is misleading because of sex differences manifest from that age upwards.
5. Mean Müller-Lyer illusion scores are lower for girls than for boys until about age 12. From ages 11 to 17 the illusion for boys steadily decreases, whereas that for girls increases, and the two curves cross at about age 13 and again at about age 18. The most pronounced sex differences occur at about 16 and 17 years.
6. When girls, ages 16^0 to 17^6, were tested for Müller-Lyer illusion under conditions designed to produce a critical attitude, the mean score did not differ significantly from that secured under the original conditions.
7. Mean vertical-horizontal illusion scores decrease consistently from ages six to 17, and increase rather markedly from ages 17 to 19. The decrease is more gradual from ages 11 to 14, and is relatively great between ages 14 and 17.
8. Boys of all ages have a somewhat higher vertical-horizontal

illusion score than have the girls, but the general trend of the two curves is the same.

9. The curve for mean variation of Müller-Lyer illusion scores at the different age levels runs counter to the mean illusion scores up to the age of 14. Beyond this age the two curves run parallel. Variation increases consistently from ages six to 11 and reaches its highest point at age 11, the age at which the lowest mean illusion score is obtained.

10. The curve of mean variation in Müller-Lyer illusion for girls rises steadily from ages six to 11, descends consistently from ages 11 to 17, and rises slightly from ages 17 to 19. The highest mean variation occurs at age 11, the age at which the lowest mean illusion score is obtained, and the lowest mean variation is obtained at age 17, the age at which the highest mean illusion score (with the exception of age six) is obtained.

11. The curve of mean variation in Müller-Lyer illusion for boys rises steadily from ages six to 11, drops sharply from ages 11 to 14, drops slightly from ages 14 to 17, and rises slightly from ages 17 to 19. The lowest mean variation occurs at age 17, the age at which the lowest mean illusion score is obtained.

12. The curve for mean variation of vertical-horizontal illusion scores goes up markedly from ages six to eight, in striking contrast to the mean illusion scores which do not change appreciably during this period. Mean variation decreases consistently from ages eight to 17, and increases markedly again from ages 17 to 19. From ages eight to 19 the trend for mean variation is in the same direction as mean illusion score.

13. Reliability coefficients for the Müller-Lyer illusion increase greatly from ages six to seven. Beyond this age there is a very slight and statistically insignificant trend toward higher reliabilities with increasing age.

14. Reliability coefficients for the vertical-horizontal illusion do not increase appreciably with advance in age but are quite stable throughout the entire age range.

15. Reliability coefficients at all age levels, with the exception of Müller-Lyer reliability at age six, are sufficiently high for group prediction.

16. Distribution curves for Müller-Lyer illusion are positively

skewed at those ages where growth is most rapid. This may, however, be a result of unreliability of measurement.

17. The vertical-horizontal illusion is symmetrically distributed at all age levels except 10, where it is positively skewed at the one per cent level of significance, and age 15, where it is positively skewed at the five per cent level of significance. The skewness at age 10 is associated with a marked change in mean illusion score. This may, however, be a result of unreliability of measurement.

18. Correlations between Müller-Lyer illusion scores over a three-year interval, with the exception of that for age six, become consistently higher with advance in age. This suggests that reactions become more controlled with increasing age. The relatively high correlations at age six may be attributed to the small number of cases. The only statistically significant correlation is at age eight.

19. Correlations between vertical-horizontal illusion scores over a three-year interval increase markedly from ages six to seven, after which they remain about the same and are statistically significant. Apparently control of this reaction is achieved at an earlier age than is reaction to the Müller-Lyer figure.

20. *Mean differences in Müller-Lyer illusion over a three-year* period are statistically significant and are opposite in direction to *the age differences obtained in the cross-sectional study.*

21. Although all of the follow-up groups increase in mean Müller-Lyer illusion over a three-year period, the relative position of the group means remains the same and is consistent with the results of the cross-sectional study.

22. A positive increment in Müller-Lyer illusion was obtained when a group of 42 children, average age seven years and four months, was retested at an interval of from five to eight weeks after the first test. This increment was statistically significant at the one per cent level. A positive and statistically significant increment in vertical-horizontal illusion was also obtained.

23. Mean differences in vertical-horizontal illusion scores over a three-year period are in the same direction as the age differences obtained in the cross-sectional study. They are statistically insignificant at age six, significant at the five per cent level at ages seven and eight, and significant at the one per cent level at age nine.

24. All of the follow-up groups decrease in mean vertical-horizontal illusion over a *three-year period. With the exception of*

the youngest age group, the relative position of the group means remains the same for both tests.

25. There is a low positive correlation between Müller-Lyer and vertical-horizontal illusion at all age levels. There is a slight tendency for the correlations to be higher at the lower age levels than at the upper age levels.

26. There is reason to doubt the reliability of Müller-Lyer illusion scores based upon four trials, since scores on four trials correlated only .4322 with scores based on 20 trials for a group of 50 children whose average age was seven years and four months. This correlation was shown to be spuriously high. Scores based upon 20 trials had a split-half reliability of .9546.

27. The reliability of vertical-horizontal illusion scores based upon four trials appears to be better established than that for the Müller-Lyer illusion, since these scores correlated .9156 with scores based upon 20 trials. This correlation, like that for the Müller-Lyer data, however, is spuriously high. Scores based on 20 trials had a split-half reliability of .9721.

28. The results of this study support the Gestalt explanation of the Müller-Lyer illusion, and lend support to Luria's contention that "the simple reaction" in young children differs from the reaction of the adult in having another structure.

29. The inconsistency of the results of the cross-sectional and of the longitudinal approach, as well as the results of the five to eight week follow-up study, suggest the need for caution in making longitudinal studies of the growth of certain psychological processes.

30. The analysis of variance technique is shown to have distinct limitations in the study of growth data in general, since its basic assumption of homogeneity of variances will seldom be fulfilled.

31. A possible relationship between geometrical-optical illusions and such perceptual phenomena as thing constancy and eidetic imagery is discussed.

REFERENCES

1. Allport, G. W. The Eidetic image and the after-image. *Amer. J. Psychol.*, 1928, **40**, 418-425.
2. Anderson, J. E. Child development and the interpretation of behavior. *Science*, 1936, **83**, 245-252.
3. Aveling, F., & Hargreaves, H. L. Suggestibility with and without prestige in children. *Brit. J. Psychol.*, 1921, **12**, 53-75.
4. Bates, M. A study of the Müller-Lyer illusion with special reference to paradoxical movement and the effect of attitude. *Amer. J. Psychol.*, 1923, **34**, 46-72.
5. Binet, A. La mesure des illusions visuelles chez les enfants. *Rev. Philos.*, 1895, **40**, 11-25.
6. Bayley, N. Mental and motor development from two to twelve years. *Rev. Educ. Res.*, 1939, **9**, 18-37.
7. Boring, E. G. Is there a generalized psychometric function? *Amer. J. Psychol.*, 1924, **35**, 75-78.
8. ———. The logic of the normal law of error in mental measurement. *Amer. J. Psychol.*, 1920, **31**, 1-33.
9. Boynton, B. The physical growth of girls: A study of the rhythm of physical growth from anthropometric measurements on girls between birth and eighteen years. *Stud. Child Wel.*, 1936, **1**, No. 4.
10. Bradway, K. P., & Hoffeditz, E. L. The basis for the personal constant. *J. Educ. Psychol.*, 1937, **28**, 501-513.
11. Bryan, W. L. On the development of voluntary motor ability. *Amer. J. Psychol.*, 1892, **5**, 123-204.
12. Buck, A. F. Observations on the observation of vertical as compared with horizontal lines. University of Chicago, *Contrib. Philos.*, 1899, **2**, No. 2.
13. Cramaussel, E. Experiences au jardin d'enfants. *J. de Psychol.*, 1927, **24**, 701-718.
14. Crosland, H. H., Taylor, H. R., & Newsom, S. J. Practice and improvability in the Müller-Lyer illusion in relation to intelligence. *J. Gen. Psychol.*, 1929, **2**, 290-304.
15. ———. Intelligence and susceptibility to the Müller-Lyer illusion. *J. Exper. Psychol.*, 1927, **10**, 40-51.
16. Culler, B. Studies in psychometric theory: XIII. The normal law; its place in quantitative psychology. *J. Exper. Psychol.*, 1926, **9**, 271-298.
17. Dearborn, W. F. The mental and physical development of public school children. *Sch. & Soc.*, 1935, **41**, 585-593.
18. Fisher, R. A. The Design of Experiments. Edinburgh: Oliver & Boyd, 1935.
19. ———. Statistical Methods for Research Workers. Edinburgh: Oliver & Boyd, 1936.
20. Fritz, M. F. Experimental evidence in support of Professor Thurstone's criticism of the phi-gamma hypothesis. *J. Gen. Psychol.*, 1930, **4**, 346-352.

21. GARRETT, H. E. Differentiable mental traits. *Psychol. Rec.*, 1938, **2**, 259-298.

22. GILBERT, J. A. Researches on the mental and physical development of school children. *Stud. Yale Psychol. Lab.*, 1894, **2**, 40-100.

23. GOODENOUGH, F. L. The development of the reactive process from early childhood to maturity. *J. Exper. Psychol.*, 1935, **18**, 431-450.

24. GOULDEN, C. H. Methods of Statistical Analysis. New York: Wiley, 1939.

25. GRIFFING, H. On the development of visual perception and attention. *Amer. J. Psychol.*, 1895, **7**, 227-236.

26. GUILFORD, J. P. A generalized psychophysical law. *Psychol. Rev.*, 1932, **39**, 73-85.

27. ———. Psychometric Methods. New York: McGraw-Hill, 1936.

28. HARTMAN, G. W., & TRICHE, A. Differential susceptibility of children and adults to standard illusions. *J. Gen. Psychol.*, 1933, **42**, 493-498.

29. HEINIS, H. A personal constant. *J. Educ. Psychol.*, 1926, **17**, 163-186.

30. HELMHOLTZ, H. Treatise on Physiological Optics. (Translated by J. P. C. Southall.) Vol. III. New York: Optical Society of America, 1925. (Pp. 234-242; 188-204.)

31. HICKS, G. D., & RIVER, W. H. R. The illusion of compared horizontal and vertical lines. *Brit. J. Psychol.*, 1906, **2**, 243-260.

32. HILDEN, A. H. A comparative study of the intelligence quotient and the Heinis personal constant. *J. Appl. Psychol.*, 1933, **17**, 355-375.

33. JAENSCH, E. R. Eidetic Imagery. (Translated from the second edition by O. Oeser.) New York: Harcourt, Brace, 1930.

34. JASTROW, J. A study of Zollner's figures and other related figures. *Amer. J. Psychol.*, 1891, **4**, 381-398.

35. JONES, F. A Study of the Ability of Preschool Children to Equate Length of Lines on the Galton Bar. Unpublished M. A. Thesis, University of Minnesota, 1931.

36. JONES, H. E. Relationships in physical and mental development. *Rev. Educ. Res.*, 1939, **9**, 91-102.

37. JUDD, C. H. The Müller-Lyer illusion. *Psychol. Rev.*, 1898, **5**, 286-294.

38. ———. An optical illusion. *Psychol. Rev.*, 1898, **5**, 286-294.

39. ———. A study of geometrical illusions. *Psychol. Rev.*, 1899, **6**, 241-261.

40. ———. Practice and its effects on the perception of illusions. *Psychol. Rev.*, 1902, **9**, 27-39.

41. KELLEY, T. L. Principles and techniques of mental measurement. *Amer. J. Psychol.*, 1923, **44**, 408-432.

42. KLÜVER, H. Eidetic imagery. In *Handbook of Child Psychology*. Worcester, Mass.: Clark Univ. Press, 1933.

43. KOFFKA, K. The Growth of the Mind. (Translated by R. M. Ogden.) New York: Harcourt, Brace, 1924.

44. KUHLMANN, F., & ANDERSON, ROSE G. Intelligence Tests for Ages Six to Maturity. Instruction Manual, Fourth Edition. Minneapolis: Educational Test Bureau, 1933.

45. KÜLPE, O. Outlines of Psychology. (Translated by E. B. Titchener.) New York: Macmillan, 1909.

46. LEWIS, E. O. Confluxion and contrast effects in the Müller-Lyer illusion. *Brit. J. Psychol.*, 1909, **3**, 21-41.

47. ———. The effect of practice on the perception of the Müller-Lyer illusion. *Brit. J. Psychol.*, 1906, **2**, 294-306.

48. LINDER, F. E. The accuracy of the constant method. *Amer. J. Psychol.*, 1935, **47**, 504-508.

49. LINDQUIST, E. F. Statistical Analysis in Educational Research. New York: Houghton-Mifflin, 1940.

50. LINE, W. The growth of visual perception in children. *Brit. J. Psychol., Monog. Suppl.*, 1931, **5**. Pp. 148.

51. LIPPS, T. Raumästhetik und geometrisch-optische Täuschungen. (Author complete reference.)

52. LOCKE, N. M. Perception and intelligence: Their phylogenetic relation. *Psychol. Rev.*, 1938, **45**, 335-345.

53. LUCKIESH, M. Visual Illusions; Their Causes, Characteristics, and Applications. New York: D. Van Nostrand, 1922.

54. LUFKIN, H. M. The best fitting frequency function for Urban's lifted weight results. *Amer. J. Psychol.*, 1928, **40**, 75-82.

55. MEENES, M. The incidence of eidetic imagery in Negro school children. *J. Negro Educ.*, 1936, **6**, 611-616.

56. MÜLLER-LYER, F. C. Optische Urteilstäuschungen. *Arch. f. Physiol.*, Suppl., 1889.

57. ———. Über Kontrast und Konfluxion. *Z. f. Psychol. u. Physiol. Sinnesor.*, 1896, **10**, 421-431.

58. NAYER, P. P. N. An investigation into the application of Neyman and Pearson's L_1 test, with tables of percentage limits. *Stat. Res. Mem.*, 1936, **1**, 38-51.

59. OESTERREICH, T. K. Zum Problem der geometrisch-optischen Täuschungen. *Z. f. Psychol. u. Physiol. Sinnesor.*, 1928, **105**, 371-385.

60. PALMER, C. E., KAWAKAMI, R., & REED, L. J. Anthropometric studies in individual growth: II. Age, weight, and rate of growth in weight, elementary school children. *Child Devel.*, 1937, **8**, 47-61.

61. PARSONS, J. H. An Introduction to the Theory of Perception. New York: Macmillan, 1927.

62. PECK, L., & WALLING, R. A preliminary study of the eidetic imagery of preschool children. *J. Genet. Psychol.*, 1935, **47**, 168-192.

63. PECK, L., & HODGES, A. B. A study of the eidetic imagery of young Negro children. *J. Negro Educ.*, 1937, **6**, 601-610.

64. PIERON, H. L'illusion de Müller-Lyer et son double mecanisme. *Rev. Philos.*, 1911, **71**, 245-284.

65. PINTNER, R., & ANDERSON, M. M. The Müller-Lyer illusion with children and adults. *J. Exper. Psychol.*, 1916, **1**, 200-210.

66. RIVERS, W. H. R. Observations on the senses of the Todas. *Brit. J. Psychol.*, 1904-1905, **1**, 321-396.

67. RUSH, G. P. Visual grouping in relation to age. *Arch. of Psychol.*, 1937, No. 217.

68. SEASHORE, C. E., CARTER, E. A., FARNUM, E. C., & SIES, R. W. The effect of practice on normal illusions. *Psychol. Rev. Monog. Suppl.*, 1908, 9, 103-148.

69. SEASHORE, C. E., & WILLIAMS, M. C. An illusion of length. *Psychol. Rev.*, 1900, 7, 592-599.

70. SHEEHAN, M. R. A study of individual consistency in phenomenal constancy. *Arch. of Psychol.*, 1938, No. 222.

71. SHUTTLEWORTH, F. K. The Physical and Mental Growth of Girls and Boys Age Six to Nineteen in Relation to Age at Maximum Growth. *Monog. Soc. Res. Child. Devel.*, 1939, 3, No. 3.

72. ———. Sexual Maturation and the Physical Growth of Girls Age Six to Nineteen. *Monog. Soc. Res. Child Devel.*, 1937.

73. SIMMONS, K., & TODD, T. W. Growth of well children: Analysis of stature and weight, 3 months to 13 years. *Growth*, 1938, 2, 93-134.

74. SMITH, W. G. A study of some correlations of the Müller-Lyer visual illusion and applied phenomena. *Brit. J. Psychol.*, 1906, 2, 16-51.

75. SNEDICOR, G. W. Statistical Methods. Ames, Iowa: Collegiate Press, 1938.

76. STRATTON, G. M. Symmetry, linear illusions, and the movements of the eye. *Psychol. Rev.*, 1906, 13, 81-86.

77. TEASDALE, H. A quantitative study of eidetic imagery. *Brit. J. Psychol.*, 1934, 4, 56-74.

78. THIERY, A. Über geometrisch-optische Täuschungen. *Philos. Stud.*, 1895, No. 11; 1896, No. 12.

79. THORNDIKE, E. L. Educational Psychology (Vol. III). New York: Columbia Univ. 1914.

80. THOULESS, R. H. The effects of errors of measurement on correlation coefficients. *Brit. J. Psychol.*, 1939, 29, 383-403.

81. THURSTONE, L. L. Psychophysical analysis. *Amer. J. Psychol.*, 1927, 38, 368-389.

82. ———. Phi-gamma hypothesis. *J. Exper. Psychol.*, 1928, 11, 293-305.

83. TINKER, M. A. Susceptibility to optical illusions: Specific or general. *J. Exper. Psychol.*, 1938, 22, 593-598.

84. TITCHENER, E. B. A Textbook of Psychology. New York: Macmillan, 1915.

85. URBAN, F. M. The method of constant stimuli and its generalization. *Psychol. Rev.*, 1910, 17, 229-259.

86. VALENTINE, C. W. Psychological theories of the horizontal-vertical illusion. *Brit. J. Psychol.*, 1912, 5, 8-35.

87. VAN BIERVLIET, J. J. Nouvelles mesures des illusions visuelles chez les adultes et les enfants. *Rev. Philos.*, 1896, 41, 169-181.

88. VERNON, M. D. Visual Perception. Cambridge: Cambridge Univ. Press, 1937.

89. WARDEN, C. J. & BAAR, J. The Müller-Lyer illusion in the ring dove, *Turtor risorius*. *J. Comp. Psychol.*, 1929, 9, 275-292.

90. WILLIAMS, M. C. Normal illusions in representative geometrical forms. *Univ. Iowa Stud. Psychol.*, 1902, **3**, 38-139.

91. WINCH, W. H. The vertical-horizontal illusion in school children. *Brit. J. Psychol.*, 1906, **2**, 220-225.

92. WINSLOW, C. N. Visual illusions in the chick. *Arch. of Psychol.*, 1933, No. 153.

93. WUNDT, W. Outlines of Psychology. (Translated by C. H. Judd.) Chicago: W. Engelmann, 1902.

94. ———. Die Geometrisch-Optischen Täuschungen. Leipsig: Teubner, 1898. Pp. 178.

Published as a separate and in *Genetic Psychology Monographs*, 1942, **25**, 157-215.

INTERPRETATION OF BEHAVIOR-RATINGS IN TERMS OF FAVORABLE AND UNFAVORABLE DEVIATIONS: A STUDY OF SCORES FROM THE READ-CONRAD BEHAVIOR INVENTORY*

KATHERINE H. READ AND HERBERT S. CONRAD

Institute of Child Welfare, University of California

*Accepted for publication by Harold E. Jones of the Editorial Board, received in the Editorial Office on October 16, 1941, and published at Provincetown, Massachusetts.

ACKNOWLEDGMENTS

It is a pleasure to express appreciation to those who have read and criticized the manuscript of the present report: to Dr. Harriet E. O'Shea, of Purdue University; to our colleagues at the Institute of Child Welfare, Dr. Harold E. Jones, Dr. R. Nevitt Sanford, and Dr. Catherine Landreth; to Dean Frank N. Freeman, of the School of Education of the University of California; and to Mr. Hubert C. Armstrong, of the Oakland (Calif.) Public Schools. The data for the present study were collected at the Purdue University Nursery School, with the counsel of Dr. O'Shea.

Institute of Child Welfare
University of California
Berkeley, California

KATHERINE H. READ
HERBERT S. CONRAD

I. INTRODUCTION

An unnecessary hiatus exists, both in understanding and appreciation, between (*a*) those who favor the "clinical" approach to the study of personality, and (*b*) those who favor the "measurement" approach. One group points to the general subjectivity and limited verification of conclusions from the "clinical" approach. The other group, with equal justice, points to the incompleteness of any uniform set of formal measurements of personality; to the general superficiality of objective measures of personality taken at face value; and, consequently, to the genuine need for subjective interpretation. No specially compromising spirit should be needed to observe that advantages and disadvantages inhere in each of the two approaches; that neither approach necessarily excludes the other; and that, in fact, measurement without clinical interpretation may be useless, while clinical interpretation without measurement (of one sort or another) is baseless.

The present study undertakes to employ both formal measurements and interpretation. An attempt is made to lessen the objection of incompleteness, by making use of ratings in all 67 traits of the Read-Conrad *Inventory* (4); our particular concern in the present paper is with four summary-measurements derived from the ratings on these traits. An attempt is made to lessen the charge of subjectivity, by restraining interpretation to a fairly simple level. That such a double-aspect procedure may combine the advantages of both the "clinical" and the "measurement" approaches, without the disadvantages of either, would of course be far too much to expect. The procedure is, rather, presented as simply one additional device in the storehouse of techniques for the study of various aspects of personality. From the data of the present study, some encouragement may be gained for tentative practical application and further investigation.

II. PROBLEM

The present study aims primarily to devise, investigate, and apply four personality-measures based on nursery-school teachers' ratings of traits in the Read-Conrad Inventory (4):

(1) Score F is a measure of desirable or *favorable* deviations of the child's behavior from the group-average.

(2) Score U is a measure of undesirable or *unfavorable* deviations of the child's behavior from the group-average.

(3) Score N is a *net* measure of the favorableness or unfavorableness of the child's behavior.

(4) Score T is a measure of the child's *total deviations* (whether favorable or unfavorable) from the group-average.

Details concerning the calculation of the various scores are given in a later section. The present report will also undertake to consider the special methodological principles involved in the scores, and the uses to which the scores may properly be put.

III. THE SAMPLE AND THE JUDGES

The ratings of behavior were made for three groups of nursery-school children, constituting 31 cases. Group *A* consisted of eight four-year-olds, rated by six judges; Group *B*, of 12 three- and four-year-olds, rated by four judges; Group *C*, of 11 two- and three-year-olds, rated by seven judges. The children came almost entirely from professional and business families. Most of the parents had had more than a high school education.

One judge rated all three groups of children; two other judges rated two groups; the total number of different judges is 13. Of these, two were staff members in the nursery school, while 11 were graduate-student assistants. All of the judges had, through preliminary practice, gained a working understanding of the instructions and trait-definitions of the Inventory (4). The ratings were made independently by each judge, at the end of a semester, after each judge had known the children for one semester or longer.

IV. THE RATING SCALE

The Read-Conrad *Behavior Inventory for Nursery School Children* (4) consists of 67 selected items of behavior, to be rated on a seven-point scale from most to least. The traits of the Inventory represent more or less standard categories of behavior, concerning which one generally wants to have information when studying any young child. For each trait, the two extremes (Levels 1 and 7) and the mid-point (Level 4) have been defined, as in the following example (*Popularity*):

> 1. Child is a favorite with the other children: other children like especially to have the child as a playmate, or as a member in group activities.
>
> 4. Average popularity; better liked by some children than by others.
>
> 7. Child is unpopular; other children seldom choose him as a playmate, and do not care to have him as a member in group activities.

A child judged as falling at some level between 1 and 4, or between 4 and 7, is given the appropriate intermediate rating. The traits of the Inventory are listed, by title only, in the next section of this paper; for a fuller description, and a discussion of the basis for selection of the 67 traits, reference is made to a previous publication (4).

V. PROCEDURE

A. Individual and Group Averages for Each Trait

For each child, the mean of his ratings by the various nursery-school teachers on each of the 67 traits of the Read-Conrad *Inventory* (4) was calculated. Similarly, for each Group (*A*, *B*, and *C*, respectively), the average[1] rating by the various nursery-school teachers on each of the 67 traits of the Inventory was calculated. Our procedure has been to *compare a child's average rating in each trait with his group's average for that trait.* Thus, the average rating is Popularity for Case G.W. (Group *A*) is 3.4; this is compared with his group's average rating in this trait of 3.9.

The use of separate averages for Groups *A*, *B*, and *C*, respectively, deserves some consideration. One advantage of this procedure is that a child's behavior is considered with regard to the group of which he is a member. Thus, in a group characterized by great friendliness (Trait No. 3), an unfriendly child would be more exceptional (really more unfriendly) than such a child in a group characterized by reserve or indifference. The present system of scoring take account of this fact. A further consideration in the present study is that Groups *A*, *B*, and *C* differ considerably in age (cf. Section III); and while, according to the directions, "each child should be rated only *in comparison with others of the same chronological age*" (4, p. 471), this direction is more easily prescribed than executed. Experimental segregation of children into groups reasonably homogeneous in age undoubtedly offers an advantage in this respect.

On the other hand, the use of separate group-averages, when the groups are small, involves the risk of excessive sampling-fluctuations. We have tried to estimate the importance of this factor by two methods. (*a*) The mean difference, without regard to sign, between the group-average for each trait in Group *A* and in Group *B* is .36; in Group *A* vs. Group *C*, is .37; in Group *B* vs. Group *C*, is .23. The corresponding "constant differences" (i.e., mean differences with regard to sign) are —.14, —.22, and —.08, respectively. Such differences, ranging from less than a tenth of a class-interval (—.08) to slightly over a third (.37) may be considered reassuring. (*b*) Another index of agreement among trait-means in the different Groups is provided by the coefficient of correlation. The correlation

[1]For Groups *B* and *C*, the mean was employed; for Group *A* (in order to avoid excessive influence by one extreme case), the median.

between trait-means in Group *A* and Group *B* is .48; between Group *A* and Group *C*, is .40; between Group *B* and Group *C* is .64. Factors depressing these correlations include not only sampling-fluctuations in the individual trait-means, but also the definitely limited variability of the trait-means (cf. Table 1 below), and possibly also real differences associated with age-differences in the three groups.

TABLE 1
DISTRIBUTION OF MEANS OF RATINGS IN 67 TRAITS, FOR GROUPS *A*, *B*, AND *C*

Mean trait-rating	Group *A* f	Group *B* f	Group *C* f
5.0-5.1	—	2	—
4.8-4.9	—	—	1
4.6-4.7	2	1	1
4.4-4.5	4	4	6
4.2-4.3	6	7	6
4.0-4.1	11	13	20
3.8-3.9	12	16	15
3.6-3.7	13	7	16
3.4-3.5	5	11	2
3.2-3.3	7	5	—
3.0-3.1	5	1	—
2.8-2.9	2	—	—
Mean*	3.75	3.89	3.97

*Computed from ungrouped data.

To the extent that sampling errors occur in the averages of traits in Groups *A*, *B*, and *C*, some effect may be expected (*a*) on the standard deviations of scores *F*, *U*, *N*, and *T* for the total sample (increasing the *SD*'s); (*b*) on the reliability of scores (the effect may be positive or negative, depending on the direction and magnitude of errors); and (*c*) on the intercorrelations between scores. The situation is undoubtedly mitigated by the fact that each child's scores are based on ratings by from 4 to 7 judges on 67 traits—so that the possibility for "cancellation of errors" (both in sampling and measurement) is greater than usual.[3] On the whole, then, it would appear that the statistical findings of the present study may be accepted without any unusual reservation.

[3]As an illustration, Group *A* might be high in, say, 20 of the 67 traits, but compensatingly low in 30 others (perfect "cancellation of errors" is of course not assumed).

B. Establishment of Direction and Extent of Individual Deviations

The 67 traits of the Read-Conrad *Inventory* (4) were classified into three groups as follows:

Class 1: those traits in which the most desirable behavior was considered to fall at level *"1"* of the seven-point scale[8] (indicating a *high* degree of the trait). The Inventory numbers and titles of the 24 traits in this group are:

1. Popularity.
2. Talking to others (social participation).
3. Friendliness to other children.
4. Friendliness to adults.
5. Social inclinations toward individual children.
6. Social inclinations toward a group.
7. Companions in activities (social participation).
8. Sense of responsibility for the group of which he is a part.
9. Sense of responsibility in play and social reactions.
10. Group coöperation.
11. Self-reliance with respect to adults.
12. Self-sufficiency.
13. Number of interests.
14. Curiosity.
15. Sustained effort in an occupation.
16. Active persistence of interests or occupation.
17. Ambition.
18. Speed of decision.
19. Finality of decision.
20. Behavior response to difficulty.
21. Recovery from emotional disturbance.
22. Adjustability to new situations.
23. Enthusiasm.
24. Happiness, cheerfulness; good-humor, agreeableness.

If, for example, a child's average rating in *Popularity* is 2.4, while the group-average for this trait is 3.9, then the child's rating differs

[8]It is possible that the "most desirable behavior" for traits of Class 1 falls not exactly at *"1,"* but somewhere between *"1"* and *"2"* (differing somewhat for each particular trait). A highly precise allocation of the point of optimum is not necessary for ordinary purposes; and has not been attempted in the present study, either for traits of Class 1, 2, or 3 (see below). With regard to individual differences in the point of optimum for different traits, cf. Section VII, *A*, 2 and Section VII, *A*, 3.

from the group-average by 1.5; and since the child's average rating is nearer the optimal rating of *"1"* than is the group-average, we count the deviation of 1.5 as *favorable*. Had the child's average rating for the trait been 5.4, his deviation from the group-average would be 1.5, *unfavorable*.

Class 2: those traits in which the most desirable behavior was considered as falling in the region about the group-average for the trait. The Read-Conrad *Inventory* numbers and titles of the 15 traits in this group are:

25. Desire for affection.
26. Affectionateness to adults and to children.
27. Behavior reaction to sympathy or approval.
28. Compliance with respect to suggestions from adults.
29. Compliance with respect to suggestions from children.
30. Compliance in regime.
31. Reaction to social pressure.
32. Reaction to teasing.
33. Rights (self assertion).
34. Attempts at leadership in group.
35. General statement of inhibition of emotions.
36. Ease with which the child is emotionally affected.
37. Degree of emotional response.
38. Restlessness.
39. Inspection of hurts.

For the traits in Class 2, all deviations from the group-average are counted as unfavorable; except that a *leeway or tolerance of half an interval is allowed* before any deviation is counted at all. If, for example, a child received an average rating of (say) 3.4 or 4.4 for a trait in Class 2 whose group-average is 3.9, no unfavorable deviation would be counted. If the child's rating for this trait were 2.4 or 5.4, his unfavorable deviation (taken from 3.4 or 4.4, respectively) would be 1.0. The traits of Class 2 reflect the fact—often optimistically overlooked—that when the region about the group-average represents the optimum, an individual's liability to unfavorable deviation is not matched by a corresponding opportunity for favorable distinction.

Class 3: those traits in which the most desirable behavior was considered to fall at level *"7"* of the seven-point scale (indicating absence of the trait). The Read-Conrad *Inventory* numbers and titles of the 28 traits in this group are as follows:

40. Dependence on outside suggestion and direction.
41. Suspiciousness of people.
42. Apprehensiveness.
43. Fear in response to bullying, threatening commands, etc.
44. Fear in non-social situations.
45. Elaboration or indirectness of response to fear.
46. Elaboration or indirectness of response to failure or frustration.
47. Emotional reaction to defeat.
48. Response when thwarted by other children.
49. Response when thwarted by an adult.
50. Bossiness in a group or with individual playmates.
51. Boastfulness.
52. Fault finding.
53. Blaming of others.
54. Bullying.
55. Ruthlessness, lack of kindness and consideration for others.
56. Irritability.
57. Quarrelsomeness.
58. Ease of stimulation of anger (other than anger displayed in temper tantrums, when child is dealing with adults).
59. Attacking others.
60. Delayed retaliation (revenge).
61. Circuitous ("compensatory") retaliation, continuing and largely indiscriminate.
62. Proneness to antipathies (intense, lasting hatred).
63. Jealousy of partiality to another child.
64. Nervous habits.
65. Stammering.
66. Special peculiarity of child's usual facial expression.
67. Sulking.

If, for example, a child's average rating in *"Suspiciousness of people"* is 2.6, while the group-average for this trait is 4.1, then the child's rating differs from the group-average by 1.5; and since the child's average rating is farther from the optimal rating of *"7"* than is the group-average, we count the deviation of 1.5 as *unfavorable*. Had the child's average rating for the trait been 5.6, his deviation from the group-average would be 1.5, *favorable*.

While there seems little room for doubt as to the general validity of the threefold classification of traits given above, it is admitted that the classification may be regarded as lacking in refinement.

It is possible for example, that a segment other than ±.5 about the group-average should have been selected as "ideal" for certain traits in Class 2. Further discussion of the general procedure is reserved for a later section of this paper (Section VII).

C. Scores F, U, N, and T

The aim of the present study is, first of all, to devise four personality measures which should be of value and convenience in the study of nursery school children. The basis of these measures is the amount by which each child's average rating in each trait deviates from the group-average (*a*) in a favorable direction, or (*b*) in an unfavorable direction (cf. Section V, *B* above). In these terms, the four personality-measures for each child are defined as follows:

1. The child's score for *favorable* behavior (symbolized F) is the sum of his favorable deviations.

2. The child's score for *unfavorable* behavior (symbolized U) is the sum of his unfavorable deviations.

3. The child's *net behavior score* (symbolized N) is the algebraic difference between his scores for favorable and for unfavorable behavior. A positive N-score denotes a net favorable balance between F and U; a negative N-score denotes an unfavorable balance.

4. The child's total-deviation score[4] (symbolized T) is the sum, without regard to sign, of his scores for favorable and for unfavorable behavior.

These personality measures are designed to indicate (*a*) the *extent* of a child's deviations from the central tendency of his group (score T); (*b*) the *direction* of the child's deviations (scores F, U, and especially N); and (*c*) the *variability in direction* of the child's deviations (score F vs. score U). The measures, it must be emphasized, are offered *for their own value,* and *not as a substitute* either for the detailed ratings on individual traits, nor for other information. Data and considerations pertinent to the validity and usefulness of the scores are presented in the following sections.

D. Measures of Relationship

Correlation in the present paper is measured by the standard

[4]The term "total deviation" is somewhat elliptical, since the "total deviation score" does not, for traits of Class 2, include deviations within ±.5 from the group-average (cf. Section V, *B* above). The term seems, however, sufficiently accurate for all practical purposes.

Pearson r, and also by the Bernstein r (1). The difference between the Pearson and Bernstein r's is analogous to that between the mean and the median, or between the mean deviation and the standard deviation. The Bernstein r is a first-power statistic, treating scores exclusively as simple arithmetic deviations from the mean; the Pearson r is, in a sense, a second-power statistic, since it is based on sigma scores, which in turn require use of the standard deviation (involving squared deviations from the mean). For a strictly normal, linear correlation surface, the Bernstein and Pearson r's are numerically identical; for a skewed or irregular correlation surface (especially one including extreme or exceptional cases), the two r's are likely to differ. In general, the extremely exceptional case (such as Case 4 in Table 2) will have less influence in determining the Bernstein r than the Pearson r. If, for example, the extreme case falls on, or very close to, the "line of relation" (i.e., the line midway between the two regression lines), such a case will add more to the magnitude of the Pearson r than the Bernstein r: the Pearson r will be higher. If, on the other hand, the extreme case is discordant, i.e., falls a considerable distance from the line of relation, such a case will reduce the Pearson r more than the Bernstein r: the Pearson r will be lower.

The Pearson r is ordinarily to be preferred to the Bernstein r, on several grounds. (*a*) The Pearson r is much the more familiar. (*b*) The Pearson r has been thoroughly investigated, so that formulas for manipulation and correction (as in the Spearman-Brown prophecy formula, and correction for attenuation) are available, and have been tested by experience. The applicability of such formulas to the Bernstein r, under conditions of non-normality, is not obvious. Finally, (*c*) the Pearson r is always definite; while the Bernstein r, being the geometric mean of two separate coefficients, r_1 and r_2 (which may differ in sign), becomes imaginary if one of these coefficients is positive and the other negative. (Bernstein $r = \sqrt{r_1 r_2}$.) This difficulty (arising only when the relation between X and Y is low and irregular) may be readily circumvented by taking the simple arithmetic mean of r_1 and r_2, instead of the geometric mean.

Although the Pearson r is, as stated, ordinarily to be preferred to the Bernstein r, we have employed the Bernstein r along with the Pearson r, because the scatter-diagrams of the present study suggested the general desirability of such a supplementary measure. For the present data, the Bernstein r's present a picture which is

generally less complimentary, though probably on the whole truer, than that given by the Pearson r's. Statistical manipulations (such as represented by the Spearman-Brown formula) will be restricted to the Pearson r; as indicated above, this is necessary because of the doubtful validity of conventional formulas when applied to the Bernstein r under conditions such as those of the present study.

VI. STATISTICAL RESULTS

A. Original Data and Statistical Constants

Lack of space prohibits a presentation of the ratings by each teacher for each child on each of the 67 traits; the resultant scores for each child in *F*, *U*, *N*, and *T* are, however, listed in Table 2.

TABLE 2
Data for Individual Cases

Case no.*	Sex	Age	Score for Favorable deviations (*F*)	Unfavorable deviations (*U*)	Net deviation (*N*)	Total deviation (*T*)
1	M	4- 0	65	7	58	72
2	M	4- 9	12	45	—33	57
3	F	4- 6	18	8	10	26
4	M	4- 9	7	102	—95	109
5	M	4- 9	34	4	30	39
6	F	4- 6	25	10	15	35
7	F	4- 9	14	36	—22	49
8	F	4- 6	11	22	—11	33
9	F	4- 5	5	16	—11	20
10	F	4- 9	19	16	3	35
11	M	4- 2	21	8	13	29
12	M	4- 2	14	22	— 8	37
13	M	4-10	10	29	—20	39
14	M	3- 7	8	35	—27	43
15	F	3- 7	25	26	— 1	51
16	F	3- 8	12	18	— 6	30
17	M	3- 5	48	3	45	51
18	F	3- 9	26	9	16	35
19	F	4- 1	13	58	—45	72
20	F	4- 0	22	17	5	39
21	F	3- 5	4	38	—34	42
22	M	3- 0	3	19	—16	21
23	M	2- 6	15	18	— 4	33
24	M	2-11	33	5	27	38
25	F	2-11	21	4	17	25
26	F	2-10	24	12	12	37
27	F	2- 7	9	28	—18	37
28	F	2- 1	18	8	10	25
29	M	2- 6	16	1	15	17
30	F	2- 7	8	9	— 1	17
31	M	2-11	2	32	—30	34

*Cases 1-8, 9-20, and 21-31 constitute Groups *A*, *B*, and *C*, respectively (cf. text, Section III).

Frequency distributions of the scores for the 31 cases are provided in Table 3. Means, medians, and standard deviations are given in Table 4.

TABLE 3

FREQUENCY DISTRIBUTIONS OF SCORES

Favorable deviations (F) Score	f	Unfavorable deviations (U) Score	f	Net deviation (N) Score	f	Total deviation (T) Score	f
		105-109	—			105-109	1
		100-104	1			100-104	—
		95- 99	—			95- 99	—
		90- 94	—			90- 94	—
		85- 89	—			85- 89	—
		80- 84	—			80- 84	—
		75- 79	—	50- 59	1	75- 79	—
		70- 74	—	40- 49	1	70- 74	2
65-69	1	65- 69	—	30- 39	1	65- 69	—
60-64	—	60- 64	—	20- 29	1	60- 64	—
55-59	—	55- 59	1	10- 19	8	55- 59	1
50-54	—	50- 54	—	0- 9	2	50- 54	2
45-49	1	45- 49	1	— 1- 10	5	45- 49	1
40-44	—	40- 44	—	—11- 20	5	40- 44	2
35-39	—	35- 39	3	—21- 30	3	35- 39	10
30-34	2	30- 34	1	—31- 40	2	30- 34	4
25-29	3	25- 29	3	—41- 50	1	25- 29	4
20-24	4	20- 24	2	—51- 60	—	20- 24	2
15-19	5	15- 19	6	—61- 70	—	15- 19	2
10-14	7	10- 14	2	—71- 80	—	10- 14	—
5- 9	5	5- 9	7	—81- 90	—	5- 9	—
0- 4	3	0- 4	4	—91-100	1	0- 4	—
Total	31		31		31		31

TABLE 4

STATISTICAL CONSTANTS OF SCORES

Measure	Favorable deviations (F)	Unfavorable deviations (U)	Net deviation (N)	Total deviation (T)
Mean	18.0	21.2	— 3.6	39.6
Median	15.0	16.6	— 3.5	36.3
SD	13.4	20.2	28.5	17.9

The frequency distributions of Table 3 serve to emphasize the extent of individual differences. Thus, the child highest in N is about 150 points or 5+ SD's above the child lowest in N; the most-deviating child (highest in T) has a score over six times as high as the least deviating. Surely such large numerical differences signify important personality differences for the children concerned.

Sadly enough, the largest score for unfavorable deviations (U) is a great deal higher than for favorable deviations (F). We might have expected that the 15 traits of Class 2, which contribute to U but not at all to F (cf. Section V, B above), would lead to a distribution for U having a higher mean and SD than for F; some such effect is observable (cf. Tables 3 and 4); but less than one would suppose—evidently a "lenient tendency" of the raters has at least partially intervened. In both the distribution for F and U, low scores are more common than high; i.e., the distributions are skewed. If the U-scores are thought of as negative F-scores, belonging in the same distribution as the F-scores, the resulting single distribution is relatively flat and extended, with a suggestion (probably unreliable) of bimodality.

Significant individual differences occur not only in the separate scores for F, U, N, and T, but also in the pattern or combination of scores. Thus, of the five highest scores for T in Table 2, three are based mainly on high scores in U (Cases 2, 4, and 19); one is based mainly on a high score in F (Case 1); while one is based equally on scores in both F and U (Case 15). Cases 10, 15, 20, and 23 have F-scores practically equal to their U-scores; whereas others have F-scores far in excess of U, or vice versa. Such differences in the pattern or combination of scores are not without significance in personality analysis, and suggest the importance of a proper balance of attention between any given score and its coordinate setting.

B. Relations to Chronological Age and Sex

The correlations between age, on the one hand, and scores F, U, N, and T on the other, are respectively, .08, .31, —.20, and .39. None of these correlations is reliably different from zero.[5] An interesting feature of the age-correlations is the perceptible tendency for scores to "fan out" as the higher CA's are reached. The reason for this is not definitely known. It is, however, possible that older children's behavior is more active, clear-cut, and unambiguous than younger children's; this would lead to a more generous use of extreme ratings by the judges for the older children, and thus account for their greater spread of scores.

There are no perceptible sex-differences. The median scores for

[5]For such unreliable relations, it has appeared superfluous to compute the corresponding Bernstein r's.

boys vs. girls are as follows: for F, 14.5 vs. 15.3; for U, 17.0 vs. 16.4; for N, —5.5 vs. —2.2; for T, 37.0 vs. 35.8.

C. Reliability of Scores

The reliability of scores F, U, N, and T has been ascertained first by a measure of inter-judge agreement, and second by the odd-even ("split-half") technique.

1. *Inter-Judge Agreement*

In this method, each child's average score in each variable, based on *half the judges'* ratings, was correlated with his average score based on the *remaining judges'* ratings.[6] Both Bernstein and Pearson correlations were calculated (cf. Section V, D above). Because the necessary detailed data were lacking for Group A at the time of the present analysis, the correlations are based only on the cases in Groups B and C ($n = 23$).[7] In each of these two groups, the judges were matched so far as possible on the basis of their estimated skill and experience in rating, 2 against 2 in Group B, and 4 against 3 in Group C (cf. Section III above, "The Sample and the Judges"). The correlations obtained may be symbolized as $r_{F_1F_2}$, $r_{U_1U_2}$, $r_{N_1N_2}$, and $r_{D_1D_2}$, where the subscripts "*1*" and "*2*" refer to ratings by the first and second sets of judges, respectively. Raising the Pearson r's by the Spearman-Brown formula—the Bernstein r's are not known to be adapted to this technique[8]—yields the reliability coefficients listed in Table 5 below. These reliability coefficients, ranging from .73 to .91, with the highest reliability of .91 for what

[6]In calculating the child's scores, the group-average employed as the point of reference for each trait is based on ratings by *all* the judges of the child's group (cf. Section V above, "Procedure").

[7]The median scores in F, U, N, and T for the abbreviated sample are virtually identical with those for the total group. The standard deviations in the abbreviated group are, however, smaller; namely, 10.6, 13.5, 20.3, and 11.9, respectively. The difference in SD's is caused almost wholly by the absence, from the abbreviated sample, of one case with highly aberrant scores (Case 4 of Group A, cf. Table 2).

[8]Cf. Section V, D above. In the present case, formal validity cannot be claimed for application of the Spearman-Brown formula to the Pearson r's since equality is lacking in the SD's of ratings by judges of Set 1 vs. Set 2 (3, p. 205). This inequality would be important, however, only if the ratings of Set 1 were definitely more (or less) reliable than those of Set 2—a contingency not, in our judgment, supported by any reasonable likelihood.

TABLE 5
RELIABILITY OF SCORES, BASED ON COEFFICIENTS OF INTER-JUDGE AGREEMENT

Variable	Inter-judge agreement Bernstein r	Pearson r	Reliability coefficient (Spearman-Brown)*
F	.71	.68	.81
U	.78	.79	.88
N	.92	.84	.91
T	.46	.58	.73

*Based on Pearson r's only (see text).

is perhaps the most important score (N), seem reasonably satisfactory by current standards and resourses in this field.

Statistical constants pertinent to the coefficients of Table 5 are presented in Table 6.

TABLE 6
MEANS AND SD'S OF SCORES FROM ODD-EVEN SETS OF JUDGES

Variable	*Mean*		*SD*	
	Set 1	Set 2	Set 1	Set 2
F	16.2	21.7	9.4	14.5
U	18.6	23.9	11.9	15.2
N	— 3.0	— 2.1	17.4	24.9
T	35.0	45.7	11.3	16.7

2. *Split-Half Reliability*

Each child's score based on the *odd-numbered traits* of the Inventory (4) was correlated with his score based on the *even-*

TABLE 7
RELIABILITY OF SCORES, BASED ON SPLIT-HALF CORRELATIONS

Variable	Split-half correlation Bernstein r	Pearson r	Reliability coefficient (Spearman-Brown)*
F	.920	.942	.970
U	.952	.961	.980
N	.956	.951	.975
T	.908	.952	.975

*Based on Pearson r's only (see text). The coefficients in the present table are based on the total group ($n = 31$). For comparison with Table 5, we have calculated the split-half correlations for the abbreviated sample; these (raised by the Spearman-Brown formula) are, respectively, .941, .980, .976, and .932.

TABLE 8

MEANS AND *SD*'S OF SCORES FROM ODD-EVEN HALVES OF INVENTORY

Variable	Mean Odd	Mean Even	*SD* Odd	*SD* Even
F	9.0	9.0	6.7	6.5
U	10.8	11.1	9.5	10.3
N	— 1.7	— 2.0	14.1	14.7
T	19.4	20.0	8.8	9.4

numbered traits. Again, both Bernstein and Pearson *r*'s were calculated. Raising the Pearson *r*'s by the Spearman-Brown formula yields the reliability coefficients given in Table 7 below. The high coefficients in this table doubtless reflect some boosting by the well-known "halo effect"; but to attribute the high split-half reliabilities mainly to this spurious factor would be incorrect. It will be recalled that the inter-judge reliability coefficients (Table 5), in which halo-effect plays no part, were also, in general, satisfactorily high.[9]

Statistical constants for the odd-even halves of the Inventory employed to obtain the split-half coefficients, are presented in Table 8.

D. INTERCORRELATIONS OF SCORES

1. *Correlation between Scores for Favorable and Unfavorable Deviations*

The two basic scores of the present study are those for *favorable* behavior-deviations (F) and *unfavorable* behavior-deviations (U); the other scores, N and T, are merely a subtractive or additive combination of F and U. If the correlation between F and U were highly negative (approaching —1.00), we should for all practical purposes have not two variables, but one. The actual Pearson correlation between F and U is —.46 ($n = 31$); for such an r, the index of forecasting efficiency is only 11 per cent (2, pp. 268-269). The Bernstein correlation (cf. Section V, *D* above) is larger, —.61, but still far below a level indicating close duplication between the two variables. Because of its basic importance, the scatter-diagram for r_{FU} has been included as Table 9 of the present report. From the scatter-diagram, it is clear that (except at the very extremes) a

[9]Other evidence of lack of a predominant halo effect is the absence of a high negative correlation between F and U (cf. Section VI, *D*, 1 below, particularly the closing paragraph on "Halo effect and the correlation between F and U").

TABLE 9

Correlation between Scores for F (Favorable Deviation) and U (Unfavorable Deviation)

U-Score (Unfavorable Deviation)	F-Score (Favorable Deviation) 0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	Total
100-104		1													1
95-99															
90-94															
85-89															
80-84															
75-79															
70-74															
65-69															
60-64															
55-59			1												1
50-54															
45-49			1												1
40-44															
35-39	1	1	1												3
30-34	1														1
25-29		1	1			1									3
20-24			2												2
15-19	1	1	1	2	1										6
10-14					1	1									2
5-9		1		2	1	1	1							1	7
0-4				1	1		1			1					4
Total	3	5	7	5	4	3	2			1				1	31

Pearson $r = -.46$

Bernstein $r = -.61$

given F or U score does not permit close prediction of a child's U of F score, respectively. The two scores would appear to require (as they have in the present study received) separate calculation and joint consideration—especially if the measures are to be used for the study not only of groups but also of individual children.

The basic importance of the correlation, r_{FU}, justifies some detailed consideration of the relationship between F- and U-scores.

a. Arithmetic necessity of some negative correlation between F and U. Since the number of traits judged for each child is fixed at 67 (the number in the Read-Conrad *Inventory;* and since the contribution of any particular trait to a child's F-score excludes the possibility of contribution to his U-score—it follows as an arithmetic necessity that there will be some negative correlation between F and U. A child can, for example, have a very F-score only through having been credited with large favorable deviations from many traits; in consequence, his U-score (except as this may be derived from traits of Class 2—cf. Section V, B above) must be low. Conversely, a very high U-score necessitates a very low F-score. But (as borne out in Table 9) a very low score in F implies no necessary restriction in the range of a child's U-score, or vice versa. The conclusion, then, is that while the correlation between T and F is necessarily negative, the degree of negative correlation is *not fixed nor necessarily high.*

b. Form of the Scatter-Diagram for r_{FU}. The fact that (a) a very high F-score necessarily implies a low U-score, (b) a very high U-score necessarily implies a low F-score, but (c) moderate or low F- or U-scores permit a broader range of possibilities—leads to a triangular-shaped scatter-diagram (Table 9), which is neither normal nor homoscedastic. Needless to say, the correlation coefficient for such a surface requires careful and judicious interpretation. Another irregularity is the fact that the regression lines in Table 9 are not straight; this, however, is due almost wholly to only three cases at the extremes of the F- or U-distribution;[10] for the bulk of the cases, linear regression lines would apparently provide a good fit.

c. Correction for Attenuation. A reasonable question is whether the correlation between F and U, although not excessively high nor rigidly fixed, would be much higher if corrected for attenuation. In considering this question, it should be recognized, first, that errors in F and U are not uncorrelated: an F-score excessively high due to error implies, in general, a U-score excessively low—in short, errors in F and U are *negatively* correlated. Negatively correlated errors will, in general, *expand* the true negative correlation between two

[10]With these three cases excluded, the correlation r_{FU} equals —.56 (Pearson) and —.68 (Bernstein).

variables, making the raw or obtained r more highly negative than the true r. Correcting the raw or obtained Pearson r of —.46 between F and U would, then, result in a "true" Pearson correlation of *less* than —.46 (i.e., closer to zero).[11]

d. Halo Effect and the Correlation between F and U. The sources of negative correlation between F and U may be briefly enumerated: (*a*) a real negative relationship between the sum of an individual's personality credits and his personality debits; (*b*) the arithmetic necessity of a negative correlation between scores for F and U calculated from a fixed number of traits (considered in Section *a* above); (*c*) negative correlation between errors of measurement in F and U (considered in Section *c* above); and finally (*d*) "halo effect," or the tendency to rate an individual uniformly favorably or uniformly unfavorably (as the case may be). A high degree of "halo effect" would obviously result in a high negative correlation between F and U. The fact that the correlation between F and U is no higher than it is from *all* these sources implies that the "halo effect" in the present ratings must be comparatively limited. We should judge that sources (*a*), (*b*), and (*c*) mentioned above would account for at least —.30 or —.35 of the Pearson Correlation of —.46 between F and U; if this is so, the contribution of "halo" to the r of —.46 cannot be large.

2. *Intercorrelations between Scores F and U vs. N and T*

The preceding section has indicated the need for joint consideration, rather than inverse substitution, of a child's scores for F and U. A systematic method to provide such consideration is to obtain each child's net score (by the formula $N = F - U$), and his total-

[11]The argument above requires certain theoretical qualifications. *If* the true negative correlation between F and U were very high, and *if* errors of measurement were quite large and variable, then (especially if the negative correlation between errors were only slight) the effect of errors of measurement would be not to expand, but to depress, the true negative correlation. None of these conditions, however, appears to be fulfilled in the present case. Reliability coefficients indicate that errors of measurement are not large (cf. Section VI, C above); and this implies that the true negative r between F and U cannot be very different from the obtained r, which is certainly not very high. The degree of correlation between errors is not known, but from mere arithmetic necessity (in view of the opposed nature of F- and U-scores) may be expected to be significant. It seems fair to conclude, then, as indicated in the text, that the true negative Pearson correlation between F and U is probably *less* than —.46 (i.e., closer to zero).

deviation score (by the formula $T = F + U$). One would expect the intercorrelations, r_{FN}, r_{UN}, r_{FT}, and r_{UT}, to be high, since F enters directly into N and T and U likewise enters directly into N and T (cf. Table 10). In spite of this, it is clear that

TABLE 10

INTERCORRELATIONS BETWEEN SCORES F AND U VS. N AND T

Correlation between	Bernstein r	Pearson r
Favorable behavior (F) and net behavior score (N)	.91	.80
Unfavorable behavior (U) and net behavior score (N)	—.93	—.89
Favorable behavior (F) and total-deviation score (T)	.07*	.23
Unfavorable behavior (U) and total-deviation score (T)	.64	.75

*Simple average of r_1 and r_2, instead of geometric mean (r_1 = —.06, r_2 = .19). Cf. Section V, D above.

a child's score for N or T provides more specific information than either his F- or his U-score alone. Even a correlation of —.89 (the highest of the four Pearson r's in Table 10) implies an index of forecasting efficiency of only slightly above 50 per cent; and this would be reduced, if the correlation were corrected for attenuation (see below). The calculation of separate N and T scores, then, from F and U, would appear justified. If the scores for N and T were not explicitly calculated, a crude mental approximation to N and T would no doubt be undertaken when a child's scores for F and U were under consideration.

The contrast in magnitude between r_{FT} (below .25) and r_{UT} (above .60) is noteworthy. For the total sample, a large part of this contrast appears due to the extremely high U-score of a single case in Group A (cf. Appendix, Table 11). However, a similar contrast occurs in the abbreviated sample (Groups B and C, $n = 23$); here r_{FT} is below .30 (Bernstein r, .21, Pearson r, .28) while r_{UT} is above .65 (Bernstein r, .67, Pearson r, .66). It appears in general (as previously indicated in Section VI, A) that, as between U and F, a very high score will more likely be U than F; this tends to raise the correlation r_{UT} beyond r_{FT}. A similar, but smaller, effect is observable in the difference between r_{UN} and r_{FN}.

Correction for attenuation would tend to lower the four intercorrelations, r_{FN}, r_{UN}, r_{FT}, and r_{UT}; because in each case, the correlation between errors of measurement is positive when the obtained r is positive, and negative when the obtained r is negative.

None of the surfaces for the four intercorrelations under consideration could be expected to be normal or homoscedastic. Thus, in the case of r_{FN}, a very high F-score (since it implies a very low U-score) can be associated only with a high N-score; but for lower values of F, such close correspondence between F and N does not necessarily hold (since low F-scores may or may not be accompanied by high scores in U).

As a partial check on the degree to which curvilinearity of regression may enter as a factor depressing the obtained correlations, we have computed the multiple correlation, $R_{T(FU)}$. Since $T = F + U$, this multiple correlation should (unless the zero-order r's are depressed by curvilinear regression) equal 1.00. The calculated value of $R_{T(FU)}$ is .991; this appears close enough to 1.00 to dispel any suspicion that the intercorrelations in Table 10 might be considerably higher, if regressions were perfectly linear.

3. *Correlation between Scores N and T*

The final correlation remaining for consideration is that between N and T. This correlation expresses the relation between *net* favorableness or unfavorableness of behavior, and the child's *total deviations*[12] from the average ratings for his group. This correlation, in the total sample of 31 cases, is about —.35 (Bernstein r, —.33, Pearson r, —.38). The small size of the correlation indicates comparative independence between the two variables. Here, as before, we are dealing with a non-homoscedastic correlation surface. Correction for attenuation would raise the correlation somewhat, since (as can readily be shown algebraically) errors in F and N are virtually uncorrelated.

---

[12]The reader is reminded that, while the term "total deviation" is convenient and sufficiently accurate for practical purposes, the term is somewhat elliptical, insofar as deviations within ±.5 from the group average are not counted for the traits of Class 2 (cf. Section V, C above).

VII. DISCUSSION

A. Methodological Considerations

The fundamental methodological characteristics underlying the four personality-scores of the present study appear to include: (*a*) reliance on nursery-school teachers' ratings of "traits" of young children, as the basic source of information concerning each child; (*b*) uniform interpretation of a given rating as favorable, unfavorable, or (for traits of Class 2) neutral; (*c*) reduction of the data to numerical scores; (*d*) the use of multi-scores (F, U, N, and T) based on a single body of original data; and (*e*) comparative independence of the procedure from a prerequisite need for other data (such as information on neighborhood conditions, sibling relationships, companions, etc.). These characteristics appear to deserve some brief consideration.

1. *Use of Nursery School Teachers' Ratings of "Traits" of Young Children*

By those who emphasize the specificity of behavior, "traits" have been criticized as oversimplifications of the multitudinous facts. By those who emphasize the organized character of behavior, "traits" have been criticized as superficial, confusing, and unrelated to the "real" facts of personality. These criticisms of "traits" are not merely a priori, but have a foundation in observation and case studies. On the other hand, the *utility* of traits has equally empirical justification (see, for example, the descriptions of cases below). Are these contradictions regarding the value of the trait-approach reconcilable? Probably not. For *some* cases, specificity is apparently so great that trait-ratings *are* a misleading oversimplification; for others, the relations among traits are sufficiently peculiar and significant to require penetrating analysis and insight if anything of much value is to be gained; but for still others—and perhaps most others—trait-ratings appear to serve as a convenient means of obtaining much useful, factual information. This may be especially true for normal nursery-school children, among whom the organization (or disorganization) of behavior has not reached its greatest complexity.

The use of trait ratings does not imply rejection or depreciation of other techniques of personality study or analysis; toward these, trait-ratings may often stand in an auxiliary or adjuvant relation.

2. *Uniformity of Interpretation of "Favorable" and "Unfavorable" Deviations*

The clinical mind may reel at the audacity of undertaking to classify behavior as favorable or unfavorable, without regard to the particular child in question. "What is favorable for Johnny may be just the opposite for Jimmy; and what is favorable for Johnny at one time may be quite unfavorable at another. Such individual differences cannot safely be ignored." Again, these objections are not merely a priori assertions, but rest on a foundation of observation and case study. But again, there is strong doubt (also based on observation and case study) whether the objections have universal validity. Our impression is that such clinically-minded objections apply more specifically to a *clinical* Johnny and Jimmy than to the garden variety of normal young children. For example, it seems self-evident that, generally speaking, "friendliness to other children," "group coöperation," "adjustability to new situations," etc. (all traits from Class 1) are traits which work favorably toward a normal child's adjustment; that "fault finding," "quarrelsomeness," "attacking others," "sulking," etc. (all traits of Class 3) work unfavorably toward a normal child's adjustment; and that extremes of behavior in "desire for affection," "reaction to social pressure," or "restlessness" (all traits of Class 2) also have generally undesirable effects.[13] It is, of course, true that "attacking others," while not generally desirable, may represent a step forward for an excessively shy, inhibited child. But such a situation is in the main temporary or exceptional; it represents a special circumstance (for one of the 67 traits of the Inventory), and as such, should be considered a condition upon the validity of the procedure of the present study, rather than a complete disqualification. Nothing in the present study of course, is intended to exclude judicious consideration of the developmental peculiarities of any individual child.

Two other considerations regarding the uniform, group-interpretation of "favorable" vs. "unfavorable" deserve notice. (*a*) The group-interpretation which is undertaken is simple and general. No effort is made to classify behavior in specific detail (as to whether,

[13]These statements seem true enough for normal young children in contemporary American culture, and particularly for such children in the typical nursery school. The dependence of adjustment on cultural circumstance and individual peculiarities is recognized.

for example, it reveals a "need" for aggression, self-abasement, succorance, etc., or whether it represents introversion, dominance, self-sufficiency, etc.); instead, the classification may be considered simple virtually to the point of obviousness, and general virtually to the point of vagueness. We do not mean to imply that such simplicity and generality represent assets in personality analysis—especially since personality "analysis" is rather beyond the scope of the present effort. But we do wish to emphasize that, whereas uniform group-interpretation might very likely fail for such complicated and specific analyses as are undertaken in clinical use of free association, dream-analysis, etc., a fair measure of success should attend the much less ambitious interpretation undertaken in the present study. (*b*) Another feature favoring validity of group-interpretation in the present instance has already been touched on in a preceding section. It is the fact that the subjects of the study are still quite young. Not even the most extreme psychoanalyst would assert that the personality of the neonate compares in complexity and difficulty with that of an adult. The complexities of personality are a product of maturation and accumulating differences in experience; in both these respects, normal nursery-school children present a simpler situation than adults. To the extent, of course, that the behavior of the nursery school child involves unique complications of motive and individual modes of emotional expression, a group-interpretation (even at the relatively simple and general level of favorable-unfavorable). may be inadequate. The relatively direct and uncalculating nature of the young child's behavior, however, makes such a contingency seem rather unlikely, except possibly for a limited number of traits, or for the atypical individual. In this connection, the possibility of "compensation of errors," through use of the full 67 traits of the Inventory, perhaps deserves mention.

Such case-data as are available (cf. Section VII, *B*, 1) support the working value of the scores which have been obtained. The desirability, however, of interpretive modification of the scores for some cases should be recognized. Prudence forbids the completely routine or blind application of any technique, especially in the field of personality.

3. *Reduction of Data to Numerical Scores*

Let us recognize at once that no single personality-score, nor any set of four personality-scores, can replace interpretive consideration

of the manifold aspects and details of a child's personality. Two children with the same total personality-scores may have widely different personalities, just as two children with the same weight or sitting-height may have widely different types of body-build. But this does not deny the possible value of such variables as weight, sitting-height, or scores *F*, *U*, *N*, and *T*—both for their own sake, and as a step toward more complete and integrated understanding.

The scores *F*, *U*, *N*, and *T* represent a convenient reduction and interpretation of a considerable mass of data. The practical need for some such condensation will be apparent to anyone who has ever been confronted with ratings on each of 67 traits by several teachers for each child in a nursery-school group. The question is, however, whether the process of reduction has involved what a critic might call "a fatal loss of detail," "a crippling neglect of refinement," or "a Procrustean disregard of essential individual differences." To such questions, unqualified *Yes* or *No* answers appear to us to smack of the demagogic or obsessional; and any answer that may now be undertaken, should be improved by further investigation and additional insight.

Many of the detailed objections to use of numerical personality scores seems to resolve themselves into criticisms of uniformity of procedure. With regard to the interpretation of "favorable" vs. "unfavorable," the basis for uniformity was presented in the previous section; and similar considerations apply with regard to the arithmetical procedure of obtaining each child's score in *F*, *U*, *N*, and *T*. The adequacy of this justification may, however, be better judged if we list specifically the objections to which a uniform arithmetic procedure is subject.

The fundamental objection is that a behavior-deviation of (say) 2 may not be equally important for all traits and children. Confining our attention to "normal" nursery-school children, and thereby mitigating the problem of individual differences (cf. Section 2 above), we may observe such possibilities as these:

a. An *unfavorable* deviation of (say) 2 may, in general, be more significant than a *favorable* deviation of 2—on the ground that desirable traits tend to be taken more or less for granted, while defects are more seriously regarded.

b. A single large *unfavorable* deviation of 3 may deserve *greater* weight than three minor unfavorable deviations of 1 each.

c. A single large *favorable* deviation of 3 may deserve *less* weight than three minor favorable deviations of 1 each.

d. An unfavorable deviation of 2, based primarily on *frequency* of undesirable behavior, may deserve greater weight than an unfavorable deviation of 2 based primarily on *intensity* of undesirable behavior.

e. A favorable rating of 1 based on inconsistent extremes of behavior (e.g., for a child whose friendliness averages +1 but fluctuates markedly) may deserve less weight than a favorable rating of 1 based on consistent behavior.

f. The assumed homogeneity or interchangeability of traits may not hold true. Thus, a favorable deviation of 2 on certain highly social traits may be more important than a numerically equal deviation on relatively asocial traits.

g. Within a single trait, there may be lacking the precise balance of favorable and unfavorable extremes, which the scoring system assumes. Thus, for Trait No. 18, Level "7" (extreme slowness or irresolution of decision) is clearly unfavorable; but is Level "1" (exceptional promptness of decision) *to an equal degree* favorable? For traits of Class 2 (cf. Section V, *B*), the assumption involved is that either extreme of a trait is equally unfavorable—again a difficult condition to fulfill exactly.

h. Finally, it may be objected that an ideal procedure would require not three general classes of traits (cf. Section V, *B* above), but many more. In fact, each trait should, we suppose, have its own point of optimum, and its own system of scoring—with possibly also appropriate modifications for each particular child.

Confronted with such a variety of specific possibilities, let us first of all admit that the scores of the present study are not exact or highly refined measures. An attempt at exactness and refinement might have been undertaken by use of elaborate weighting, or conditional scoring (the score for a given trait being dependent on the nature of the child's other ratings). But probably exactness and refinement cannot be won in this field without intensive case-study procedures; and these do not seem efficiently translatable into a numerical scoring system. In any event, lack of refinement does not prevent scores *F*, *U*, *N*, and *T* from serving as definitely useful first approximations of the variables they aim to assess. The very multiplicity of detailed factors mentioned above suggests that a fair degree of compensation of errors may take place; and this suggestion

is strengthened by the fact that the scores are based on ratings by several teachers on many traits.

A final objection to numerical personality scores is that they lend themselves easily to over-use or mis-use. This is, in a sense, a compliment to the convenience of numerical scores; but it may also be viewed as a disadvantage. All personality scores involve uniformities of procedure which, in one respect or another, may prove inapplicable for certain exceptional cases. Unfortunately, there is no explicit safeguard against such mis-application, except keen appreciation of the danger, and personal familiarity with the children in question.

After all has been said that can be said against personality scores, the fact remains that such scores have their advantages. The service of scores *F*, *U*, *N*, and *T* in reduction or condensation of a large mass of data has already been mentioned. The outcome of this particular condensation is a numerical measure (or first approximation) of four variables which are of self-evident significance. The convenience of such numerical scores for records and statistical studies is obvious. Granted that comprehensive understanding of the "whole personality" requires measurement of detailed personality-components, there should be no quarrel with an attempt to measure such factors as *F*, *U*, *N*, and *T* as cheaply and conveniently as possible. It should not be overlooked, moreover, that often enough a personality score merely gives systematic, quantitative expression to a mental process which—in a crude and not-fully-conscious manner—would take place anyway. Certainly an explicit numerical score is better than a hazy, inaccurate, and only half-recognized mental surmise.

It is desirable, finally, to make explicit the precaution implied at the beginning of this section; namely, that personality scores are of limited value for the study of highly complex characteristics or individuals. Even if the scores were perfectly valid, there would still be need for much additional information and interpretation. If the scores of the present study are reasonably successful, it is probably because they aim merely to measure rather general variables, which do not (for the usual nursery-school sample) require much delicate "weighing and balancing," careful modification for exceptional factors in the situation, etc. This does *not* imply lack of significance for such scores: intelligence and weight, for example, are also variables measured by relatively simple and uniform procedures, yet are of unquestioned importance.

4. *Use of Multi-Scores*

Multi-scores (or several scores obtained from a single body of original data) are generally subject to various objections. (*a*) In the first place, there may be only a limited basis, either empirical or psychological, for the particular multiple significance assigned to a given response. (*b*) Any error in the original data results in manifold (and correlated) error in the multi-scores. (*c*) The scoring-weights employed tend to fix the correlation between the multi-scores; this correlation may correspond to the intrinsic reality (i.e., the real correlation between the traits according to external criteria), or it may not. Considering these points in the order named, we may observe, with reference to the scores of this study: (1) Although the present study employs multi-scores, no trait-rating has been assigned multiple significance. We have not, for example, said that a rating of 6 on "friendliness to other children" carries meaning with regard to the child's extroversion, dominance, neuroticism, etc. The significance of any trait-rating has been limited exclusively to a value on a "favorable-unfavorable" continuum; and the decision as to whether the given trait-rating belongs in the "favorable" or "unfavorable" portion of the continuum is hardly subject to much question, for the usual nursery-school child (cf. Section VIII, *A*, 2 above). (2) A judge's error in rating any particular trait will generally affect all four of the scores F, U, N, and T; but it should be remembered that there is considerable opportunity for compensation of errors, since each score is determined not by the rating of one judge on a single trait, but by several judges on many traits. (3) The correlations among F, U, N, and T are doubtless to some extent controlled by the arithmetical procedures employed (cf. Section VI, *D* above); but because (*a*) a zero-contribution of a trait to U may be accompanied by either a *large or small* contribution to F (or vice versa); and (*b*) the traits of Class 2 contribute exclusively to U (Section V, *B*)—it follows that the correlation between F and U is allowed considerable freedom. This point is of prime importance, since (as may be shown algebraically) all other correlations among the four scores depend principally on r_{FU}.

It remains only to consider whether, in fact, the correlations among F, U, N, and T are so high as to render one or more of these variables superfluous. Data on this point have already been presented in Section VI, *D*. It is clear that if a child has a *very*

high score for favorable deviations (F), the remaining variables (U, N, and T) are all closely predictable from the single score for F. The same is true, of course, if the child has a very high score for U. But very high scores in either F or U are, ipso facto, comparatively rare. For the great majority of cases, the range of concomitant scores in F, U, N, and T is such that no one score may safely be taken as a measure of another.

5. *Comparative Independence of Procedure*

Data from an intelligence test can ordinarily be interpreted without extensive reliance on much information external to the test. Methodologically speaking, the intelligence test is a comparatively independent procedure. The opposite is true of such a technique as the free association test (as applied in the analysis of personality); here even an initial understanding or interpretation of the facts from the test is difficult, without accessory or supporting information on the history, likes, interests, etc., of the person undergoing analysis. The contrast, then, may be quite great between what may be termed independent vs. dependent techniques. Other things being equal the independent technique has obvious advantages in lower cost and greater speed. Frequently, too, the dependent technique will require information which cannot be readily obtained.

One further advantage of the independent technique deserves notice. The dependent technique is, in general, highly subjective: it typically requires shrewd insight based on broad experience; a delicate "weighing and balancing" of many facts and possibilities; with finally an interpretation which is almost inevitably personal, frequently debatable (in one aspect or another), and practically unverifiable (by the usual scientific standards of verification). Dependent techniques have their place in the study of personality; for some cases and problems, any approach less imaginative, clever, and complicated will yield practically nothing. But for individuals within the range of normality, it appears that such techniques may be more than is needed. For normal individuals, a comparatively independent technique, such as that leading to scores F, U, N, and T, may be sufficient. Further study is always desirable, of course, if funds and resources permit, or if the child presents marked idiosyncrasies. This, however, is the exception rather than the rule. To use an analogy from medicine, we need not take X-rays, open the

abdomen for examination, and call in a staff of consultants for every person receiving a regular, yearly physical inspection.

Independence of a technique, in the sense used above, does not imply that the findings obtained from the technique will have no relations to other data. A wholly independent datum, with absolutely no relation to anything else, would of course be worthless. A datum may, however, be independent in origin or initial interpretation, and yet closely interrelated with many other data. Thus, the diagnosis of syphilis or tuberculosis is relatively independently arrived at; yet such a diagnosis may help to explain many other facts about the patient, both physical and psychological. The same applies equally to results from an intelligence test and possibly the scores of the present study. Not all independent data, of course, are equally fruitful or significant. The possibility, too, that external information will offer aid for a fuller appreciation or understanding of the comparatively independent data, should not be overlooked.

B. Uses

The scores of the present study rest upon use of the Read-Conrad *Inventory* (4); but such scores could, of course, be equally well derived from any comparable Inventory similarly applied. The essentials of the method evidently include (*a*) an Inventory composed of a large number of traits, well-defined, and well-distributed over the various aspects of nursery-school children's behavior; and (*b*) ratings for each child by as many competent judges as possible (preferably not less than four or five). With these conditions fulfilled, the uses outlined below are independent of the particular Inventory that may be applied.

1. *Uses in Nursery-School Management*

One obvious use of scores F, U, N, and T is to help select exceptionally well-adjusted and exceptionally poorly adjusted children. In this process, interpretive qualification of the scores may occasionally become important; for example, while a large negative score in N may usually be taken as indicative of maladjustment, it is possible for a child with a considerably smaller N-score also to be maladjusted. As previously stated, the scores of the present study are not exact or highly refined measures. All four scores should be considered for each child; and back of these, if necessary, the judges' ratings on the individual traits of the Inventory.

It goes without saying that the poorly adjusted child in nursery school requires special attention, both for his own good and the good of the other children. But the exceptionally well-adjusted child may also stand in need of attention; in connection with such a child, it should be considered whether the nursery-school situation may not be encouraging merely pleasant stagnation. An extremely well-adjusted child in nursery school may, in the realm of personality, be in much the same situation as a child of 140 *IQ* in school required to do only work suited to the 100-*IQ* level. Satisfactory adjustment is one goal of the nursery school; but development is another; and the developing child, presented with situations sufficiently difficult to challenge his powers, is not likely to exhibit favorable characteristics exclusively.

A second use of scores *F, U, N, T*, and the trait-ratings on which they are based, is for study and review prior to interviews or consultations with parents. Here the data are of value not only for their own sake, but also, sometimes, as a means of focussing the discussion.

A third use of the scores from the Inventory may be found in the *relation of a child's F-score to his U-score;* this relation probably has value for indicating the general type of guidance needed by the child, and the degree of success likely to be achieved. Suppose, for example, that a child has a very *high unfavorable* score, and a very *low favorable* score (e.g., Case 4, Table 2): the chances are that such an extensively unfavorable adjustment implies a personality organized along definitely undesirable lines, rather than a mere aggregation of separately acquired, inappropriate reaction-units. If this is so, a procedure restricted to superficial "habit re-training" is hardly likely to yield prompt or far-reaching success. Such a child would appear to call for careful analysis by case-study procedures; a serious effort to elicit home coöperation would doubtless be necessary, with possibly even an attempt to obtain for the child a complete change of environment (at least for a limited period). If, on the other hand, a child's high *U*-score is accompanied by an equally high *F*-score (e.g., Cases 10, 15, and 20, Table 2), we should have more hope that well planned nursery-school experience might be able to swing the balance, without the assistance of relatively elaborate and radical procedures.

Not all children, of course, run a behavior-course in accord with the theoretical expectations just described; but practical experience

suggests the feasibility of tentative use of scores F, U, N, and T in the manner indicated (see *Case Reports* below).

Case Reports

Case 1 has the highest favorable-deviation (F) and the highest net-deviation (N) scores of the whole group (cf. Table 2). This child had undergone marked unfavorable changes in environment; yet the staff members who had known him for two years had seen little change in his behavior. He was friendly, social, yet self-sufficient. Emotionally he was well balanced, responsive, and not unduly disturbed by difficulties. In his play activities he was original and purposeful, carrying elaborate projects through to completion. He stood ready to investigate anything. He was well liked by the other children.

He was the first child born at the end of the second year of marriage to well-educated and apparently well-adjusted young parents. Both parents were devoted to the child, but firm and consistent in their management of him. When the child was two-and-a-half, his mother died following the birth of the second baby. For more than a year he had been cared for by an unsympathetic mother substitute, unaccustomed to young children, who apparently disliked him and discriminated against him in favor of the baby. The behavior ratings reported were made during the period when his environment was most unfavorable. Even this marked difference in environment had not served to change a remarkably consistent and apparently stable behavior pattern.

Case 4 has the highest unfavorable-deviation (U) and the lowest net-deviation (N) score of the whole group (cf. Table 2). This child had been in one other nursery school; reports from it indicated similar unfavorable behavior. The child was timid, unfriendly, destructive, bullied the other children, and seldom engaged in constructive activities. The other children played more freely when he was not present. Careful attention to his problems in nursery school and special handling failed to effect significant improvement. Radical situational changes, directed toward modifying the child's underlying conflicts and correcting his basic attitudes, would seem to be the only effective procedure for such a case. Unfortunately, the home situation was such as to render such thoroughgoing treatment entirely impractical. Prolonged lack of parental understanding of the child's needs, and the lack of

genuine interest and affection had apparently produced habits and attitudes which could not be overcome or adequately redirected by nursery-school technique.

Cases 10, 15, and *20* are examples of cases with fairly high total-deviation (T) scores derived almost equally from scores for favorable and unfavorable deviations (cf. Table 2). It is especially in such cases, as stated above, that one would hope that nursery-school training might be able to swing the balance in a more definitely favorable direction. The reports given below on Cases 10, 15, and 20 are based first on the observations and scores for the children during their membership in the nursery-school group supervised by one of the authors of this paper; and second, either on data obtained from home visits, or from observations made in an older nursery-school group or in kindergarten.

Case 10 was the only child of older parents, both of whom were professionally employed. They were proud of *C.*, and strict and careful with her. The behavior ratings of the present study were made during the child's last semester in nursery school. *C.* was a large, placid child, advanced in her language development but lacking in motor skill. At the beginning, she sought the company and attention of teachers, largely ignoring the other children in the school. As she slowly became interested in the companionship of others of her own age, she made approaches which consisted of disparaging remarks and mild aggressions. She was not popular, but her interest and skill in social relationships steadily, though slowly, improved. The kindergarten teacher who had *C.* the next semester reported her contacts with other children quite satisfactory, and in fact pronounced her an exceptionally well-adjusted child.

Case 15 was extremely shy, dependent on the teachers, and lacking in self-assertion when she entered nursery school. Gradually, with encouragement from the teacher and support from a few pleasant and successful experiences, the child began to enter into social activity. She excelled in rhythms, and soon began to lead the children in this field. By the end of the second semester in nursery school she was playing in a group, standing up for her own rights, and showing a spontaneity and breadth of play interests which had been lacking earlier. The ratings of the present study were made during the child's first semester in nursery school; ratings during the second semester would undoubtedly have been more favorable.

Her home environment had improved at about the same time that she had entered nursery school. She and an older sibling had been cared for by relatives after her mother's death, until the father remarried and re-established a home. The stepmother, although unaccustomed to children, was sympathetic and interested. The combination of better home conditions and nursery-school experience was apparently sufficient to yield sustained improvements for this child.

Case 20 has F and U scores which are about equal and fairly large (22 and 17, respectively; cf. Table 2); so that she falls into the group for whom one might hope nursery school would prove definitely beneficial. Throughout the child's semester in nursery school, however, she remained asocial and apart, making few contacts with either teachers or children. She did not stand around watching, but quietly and effectively followed her own interests. She would sit for a long time at the table, cutting and pasting. She shared equipment willingly, exhibited a kind of passive friendliness "on request"; but she tended definitely to withdraw from contacts with others. The most that can be said for the nursery school in this case is that the child's social adjustment did not grow worse; and it is, of course, speculatively possible that the semester of nursery school did to some extent weaken the child's psychological isolation from other children, even though positive behavior-changes to support this view were not observed. One adverse factor with which the nursery school had to contend was frequent absence by the child because of colds.

Another possible factor was the mother's preference for *M.* over her older sibling, apparently because *M.* was more docile and affectionate. Very likely the mother really preferred to have the child at home than in school; at any rate, the girl was not returned to nursery school the following semester.

The case reports given above illustrate the kind of evidence for the hypothesis that a child with a high F- or (U-) score will usually resist considerable adverse (or favorable) influence; and that a child with a high T-score, derived equally from F and U, may prove especially subject to change. The evidence is, of course, incomplete. Aside from the need for a much larger number of cases, attention must also be given to possible hereditary determinants (direct or indirect) of child personality. Such an extensive inquiry, however, falls far beyond the scope of the present paper.

2. *Educational Values*

At the very least, the procedures necessary to obtain the scores *F, U, N,* and *T* should encourage alert, well-rounded observation by the nursery-school teachers and their assistants. Further benefits depend in part on the morale of the nursery-school staff, and the time available for discussion and study. The ratings and scores provide a natural and convenient starting point for study of the needs and progress of individual children, for planning of programs appropriate to exceptional cases, etc. In a school where a truly democratic spirit prevails, discussion of discrepancies in ratings or scores from different teachers can lead to improved understanding of the individual child, and also to superior insight and observation by the staff. With such discussion, there is no danger that observation will become a "closed circle," limited merely to the traits of the Inventory. Finally, it is not too much to hope that the ratings, followed by free and intelligent discussion of ratings and scores, may lead to the thoughtful, questioning, research attitude which is necessary for basic progress in any field.

The educational values mentioned above apply equally, if not more so, to students-in-training. It is not sufficiently recognized how easily students—unless given proper instruction and experience—can develop peculiar "blind spots," narrow or limited viewpoints, and a simple, "rule of thumb" set of procedures. Of special value to a student is the comparison of her trait-ratings and scores for a given child, with the ratings and scores by the nursery-school teacher, and by the group of students as a whole. Competently guided discussion of such comparisons can be very instructive and helpful. In the proper educational program for students of the nursery school, it is clear that observation, ratings, and scores such as those of the present study have a definite and significant place.

3. *Research Uses*

One of the practical and scientific questions before nursery schools today is whether nursery schools exert a perceptibly favorable effect on preschool personality, and if so, how much and how long. In the study of questions like this, scores such as *F, U, N,* and *T* offer many advantages. They provide information on variables of self-evident importance. The scores (if based on ratings from a sufficient number of judges) are reasonably reliable; they are experimentally

reproducible by different investigators in different centers; they are readily obtainable at comparatively low cost; and they are statistically manageable and convenient. To be sure, such scores are not the only type of data wanted in a study of preschool personality; detailed consideration of individual traits is also desirable, as well as study and analysis by more subjective and interpretive procedures. Nevertheless, in the battery of available techniques, scores such as those of the present study may provide one of the most feasible, convenient, and convincing approaches. These scores, as we have emphasized, are *not substitutive or competitive* in relation to other devices or techniques; they may, however, safely be accorded the rank of *e pluribus unum.*

C. Suggestions for Further Research

The present study represents a first exploration into the interpretive condensation or "numerical integration" of behavior-ratings on the Read-Conrad *Inventory* (4). Now that initial results have indicated the value of the general procedure, it is proper to consider possibilities of refinement, and to present suggestions for further research. The following list, while not exhaustive, may prove serviceable for the immediate future.

1. It has been suggested that scores F, U, N, and T cannot be taken at face value for the "clinical," "exceptional," or "peculiar" individual. What proportion of cases falls in this significantly "non-normal" category; and what uses may the scores have for such cases?

2. Would scores such as those of the present study prove useful for individuals more mature (and presumably of more complex personality-organization) than nursery-school children? At what age, if any, do such scores lose their general applicability?

3. The segregation of behavior-traits into Classes 1, 2, and 3 invites verification and refinement. One approach would be to determine whether children, with e.g., highly favorable scores receive, on the average, favorable ratings on *each* of the traits of Class 1. Another approach would require solicitation of the judgments of experts. Evidence from these two methods might or might not agree. In the event of agreement it might be desirable to establish a separate point-of-optimum for each trait (instead of, e.g., simply "1" as the optimum for *all* traits of "Class 1").

4. Should a child's behavior-deviation on a given trait be based on the child's departure from the mean of his own group on the

trait, or on the mean of a *larger* "standardization sample?" There are arguments in favor of either procedure (cf. Section V, *A*), with empirical trial desirable in order to reach a validated conclusion.

5. A distinction may be made between the general or average point-of-optimum for a trait, and the *personal* optimum (i.e., the rating judged to represent the ideal score for a given child on the particular trait). Investigation is needed to determine whether judges are able (with sufficient reliability and validity) to indicate such a personal optimum for each child. If they are, then the personal optimum on each trait could be employed as the reference point in determining a child's scores in *F, U, N,* and *T*. In this connection, a useful comparison may be possible between scores based on the general vs. the personal optimum: a large discrepancy might, for example, indicate general non-conformity, and very likely also maladjustment.

6. To what extent does the general point-of-optimum for each trait vary in different cultural groups (rural vs. urban, underprivileged vs. well-to-do, negro vs. white, Mexican-American vs. native white, etc.)? To what extent does the segregation of traits into Classes 1, 2, and 3 hold good in such different groups?

7. Do the traits of Classes 1, 2, and 3 contribute equally to scores *F, U, N,* and *T?* If not, should a system of weighting be devised in order to obtain at least approximate equality? Other problems of weighting are indicated in Section VII, *A*, 3 above.

8. Could improvement of reliability and validity be better obtained by use of more judges, or by use of a longer scale than the Read-Conrad *Inventory?*

9. What are the special formulas required to correct the intercorrelations among scores *F, U, N,* and *T* for attenuation?

10. The scores of the present study are definitely not intended as a summary of a child's entire personality; nevertheless, it would be of interest to discover to what extent children receiving a similar pattern of scores in *F, U, N,* and *T* are, in fact, of similar (or different) personality make-up.

11. As usual in this field of measurement, there is need for the development of valid criteria, against which the degree of validity of the obtained scores could be tested. Case-studies (such as in Section VII, *B* above) represent a step toward a useful criterion, but are not sufficiently quantitative to serve with complete satisfaction.

12. The convenience of numerical scores sometimes leads to their undiscriminating over-use or mis-use. What specific techniques or procedures might be devised to reduce this hazard?

While the present authors feel some obligation to pursue the problems enumerated above, it would be preferable, from a scientific point of view, for further studies to issue from new and independent sources. Some of the most pressing problems, moreover, lie quite beyond the present authors' facilities or opportunities.

VIII. SUMMARY AND CONCLUSIONS

1. *Interpretive scores.* The following four scores have been devised to provide an interpretive summary of ratings for children in the 67 traits of the Read-Conrad *Inventory*:

a. Score F is a measure of desirable or *favorable* deviations of the child's behavior from the group-average.

b. Score U is a measure of undesirable or *unfavorable* deviations of the child's behavior from the group-average.

c. Score N is a *net* measure of the favorableness or unfavorableness of the child's behavior.

d. Score T is a measure of the child's *total deviations* (whether favorable or unfavorable) from the group-average.

Details concerning the calculation of these scores are presented in Section V of this paper. The scores obviously summarize only those aspects of the ratings related to (*a*) favorableness of behavior, and (*b*) extent of deviation from the group-average. It is hoped that the scores may give systematic, quantitative expression to some of the mental processes typically employed in evaluating a child's personality from ratings. The scores should also serve in an auxiliary or adjuvant relation to other types of personality data, or other modes of analysis.

2. *Sample and judges.* The sample consists of 31 children from the Purdue University Nursery School, aged from two to five years, and each rated by from four to seven judges.

3. *Individual differences in scores and pattern of scores.* Wide individual differences occur not only in the scores for each of the variables F, U, N, and T, but also in the pattern or combination of scores. Thus, a high score in T may be derived from a high F-score, a high U-score, or a moderately high F-score in conjunction with a moderately high U. Joint attention to all four variables is therefore recommended; for detailed study, attention must also be given to the individual traits of the Inventory contributing most heavily to a given child's scores.

4. *Correlations with age and sex.* No significant correlations were found between scores on the one hand, and chronological age and sex on the other.

5. *Reliability.* As measured by inter-judge agreement (raised by the Spearman-Brown formula), the reliability of score F in the present study is .81; of U is .88; of N is .91; and of T is .73 (cf. Table 5). At least for scores F, U, and N, these reliabilities appear

high enough to justify guarded or discriminating use of the scores for the study of individual cases. Reliabilities by the "split-half" techniques are considerably higher (cf. Table 7). The discrepancy between the reliability coefficients by the two different methods may doubtless be ascribed mainly to the well-known "halo effect." The reliabilities as measured by inter-judge agreement, it should be observed, are free at least from any direct "halo" influence.

6. *Intercorrelations.* The intercorrelations among the four scores have been calculated, with a view to observing the degree of overlap or duplication in the four variables (cf. Section VI, *D*). The basic intercorrelation is that between favorable and unfavorable scores (r_{FU}), since all the other intercorrelations depend principally on this. For this basic correlation, the Pearson r is —.46, the Bernstein r (cf. Section V, *D*) is —.61. Neither of these r's indicates close duplication (the index of forecasting efficiency for a Pearson r of .46 is 11 per cent). Because of the negative correlation between errors of measurement in F and U, correction of r_{FU} for attenuation would result in a "true" Pearson r lower than —.46 (i.e., closer to zero). The data indicate that, for the great majority of cases, joint consideration of the four scores is desirable.

7. *Methodological considerations.* The fundamental methodological characteristics underlying the four personality-scores of the present study include: (*a*) reliance on nursery-school teachers' ratings of "traits" of young children; (*b*) uniform interpretation of a given rating as favorable, unfavorable, or neutral (cf. Section V); (*c*) reduction of the data to numerical scores; (*d*) the use of multiscores (F, U, N, and T) based on a single body of original data; and (*e*) comparative independence of the procedure from prerequisite need for other data (such as information on neighborhood conditions, sibling relationships, companions, etc.). Considerations bearing on these characteristics are summarized briefly below.

a. Use of ratings of "traits." Probably the two prime objections to the use of "traits" relate to specificity and superficiality. It is doubtless true that for *some* cases, specificity of behavior is apparently so great that trait-ratings involve misleading oversimplification. For other cases, the relationships among traits may be so peculiar and significant as to require abstruse analysis and insight if anything of much value is to be gained. But for *most* children (especially, perhaps, among very young children, with presumably less complex

psychological organization) trait-ratings appear to serve as an efficient means of obtaining a variety of useful, factual information.

b. Uniformity of interpretation of "favorable" and "unfavorable" deviations. While it is true that "what is favorable for Johnny may be just the opposite for Jimmy," such a comment would appear to apply principally to a *clinical* Johnny and Jimmy. Generally speaking, it seems fairly self-evident that (to select just one trait) "friendliness to other children" works favorably both toward a child's adjustment and the adjustment of his playmates; while "sulking," "attacking others," etc., work unfavorably. "Attacking others" may, to be sure, represent a step forward for an excessively shy, inhibited child. But such a situation is, in the main, temporary or exceptional. It represents a special circumstance which—unless many additional traits of the Inventory are similarly involved—may suggest the desirability of occasional subjective modification of a child's scores, but can hardly justify rejection of the technique.

c. Reduction of data to numerical scores. The scores *F, U, N,* and *T* represent a reduction and interpretation of data for a child, based on ratings of 67 traits. The practical need for some such condensation is apparent; the convenience and manageability of numerical scores are also beyond dispute. There are, nevertheless, various reasonable objections to a uniform numerical procedure; such, for example, as that an *unfavorable* deviation of 2 may generally prove more significant than a favorable deviation of 2, or that a favorable deviation of 2 in a social trait may generally be more significant than an equal deviation in a non-social trait; etc. Such considerations call attention to the fact that the scores of the present study are not exact or highly refined measures. Exactness and refinement in this field are probably not possible without intensive case-study procedures. Merely as they stand, however, the scores can serve as useful first approximations of the variables they aim to assess. The basis for this statement lies in the statistical findings mentioned above, as well as in practical experience in use of the scores (cf. Section 8 below). The chief necessary precaution is to avoid routine application of the scores when dealing with an unusually complex individual or situation.

d. Use of multi-scores. The usual objections to multi-scores have only limited applicability to the scores *F, U, N,* and *T.* (*a*) No trait-rating has been assigned multiple significance. We have not, for example, said that a rating of 6 on "friendliness to other

children" carries meaning with regard to the child's extroversion, dominance, neuroticism, etc. (*b*) A judge's error in rating any trait will affect all four scores, but only to the extent that the particular error is uncompensated by ratings on the remaining 66 traits, and by the other judges. (*c*) The basic intercorrelation, namely r_{FU} is fairly low (see above).

e. Comparative independence of procedure. A procedure which requires considerable accessory or supporting information (as to the subject's history, likes, interests, etc.) may be termed dependent; the opposite type of procedure may be termed independent. In general, the independent technique will be more objective, faster, and less costly; while the dependent technique (such as the free-association procedure) will be subjective, and permit wider application of the examiner's or analyst's insight, experience, interpretation, etc. Given the usual limitations of funds and resources, it appears likely that for individuals within the broad range of normality, comparatively independent procedures (such as lead to scores *F*, *U*, *N*, and *T*) must generally be considered sufficient. Further study by additional techniques is of course desirable for the atypical case.

Independence of a technique, in the sense used above, does not necessarily imply that the findings or scores from the technique will be lacking in relations to other data.

8. *Uses.* Scores *F*, *U*, *N*, and *T* lead themselves to a variety of uses.

a. Uses in nursery school management.

(1). The scores help to select exceptionally well-adjusted and exceptionally poorly adjusted children; both these kinds of children may require special attention.

(2). The relation of a child's *F*-score to his *U*-score probably has value for indicating the general type of guidance needed by the child and the degree of success likely to be achieved by treatment. Thus, a child with a very *high unfavorable* and a very *low favorable* score suggests a warped personality organization, rather than a mere aggregation of separately acquired, inappropriate reaction-units. In such a case, superficial "habit re-training" (while possibly a proper part of treatment) could hardly be expected to yield prompt or far-reaching success. If, on the other hand, a child's high *U*-score is accompanied by an equally high *F*-score, one might have more hope that suitably planned nursery school experience should be able to swing the

balance in the desirable direction. Illustrative case reports bearing on these hypotheses are presented in the body of the present study.

(3). Finally, both the basic trait-ratings and the scores are of convenience and value for examination prior to consultations with parents.

b. Educational values. At the very least, the procedures necessary to obtain the scores F, U, N, and T should encourage alert, well-rounded observation by the nursery-school teachers and their assistants. In a school where a truly democratic spirit prevails, discussion of discrepancies in ratings or scores from different teachers can lead to improved understanding of the individual child, and also to superior insight and observation by the staff. These educational values apply equally, if not more so, to students in training.

c. Research uses. Scores F, U, N, and T have the following advantages for research: (*a*) they offer information on variables of self-evident importance; (*b*) they are reasonably reliable (especially if ratings are obtained from a sufficient number of judges); (*c*) they are experimentally reproducible by different investigators in different centers; (*d*) they are statistically manageable and convenient; and (*e*) they are readily obtainable at comparatively low cost.

9. *Suggestions for further research.* A variety of problems has been listed for further research, such as possible modifications of scoring for children in different cultural or socio-economic groups; use of a personal vs. a general point-of-optimum for each trait; the applicability of scores such as those of the present study to older age-groups; problems of weighting; etc.

Scores F, U, N, and T are not presented in wishful negation of the manifold complexities of personality, nor as a substitute for other desirable procedures. Probably no set of scores can encompass all the essential elements and interrelations organized in the pattern of an individual personality. What the scores of the present study do aim to provide is an interpretive summary, in terms of favorable and unfavorable behavior-deviations, of ratings on the 67 traits of the Read-Conrad *Inventory.* The traits of the Inventory are more or less standard categories of behavior, concerning which one generally wishes to have information when studying any young child. The classification into favorable or unfavorable, while occasionally unsuited to the exceptional or atypical case, is scarcely subject to serious question as applied to most children. On the basis of practical

experience, supported in some respects by statistical analysis, it seems safe to say that the scores can find service for such varied functions as guidance, consultation with parents, professional stimulation and growth, teacher-training, and research. The scores are, to be sure, subject to thoughtless or unintelligent mis-use; but this is a limitation from which probably no technique in the field of personality can claim to be exempt.

APPENDIX

In order to discover the extent to which the statistical findings of the present study have been influenced by a single exceptional case (Case 4 of Group *A*, cf. Table 2), we have re-done certain calculations, excluding this case. Table 11 presents coefficients for

TABLE 11

COMPARISON OF CORRELATIONS FOR SAMPLE INCLUDING CASE 4 ($n = 31$) AND FOR SAMPLE EXCLUDING CASE 4 ($n = 30$)*

Correlation	$n = 31$	$n = 30$
Reliability coefficient of F	.970	.970
Reliability coefficient of U	.980	.962
Reliability coefficient of N	.975	.962
Reliability coefficient of T	.975	.953
$F \times U$	—.46	—.51
$F \times N$	.80	.87
$F \times T$	.23	.47
$U \times N$	—.89	—.84
$U \times T$	.75	.50
$N \times T$	—.38	.02

*The correlations for the sample of 31 cases are taken from Table 7 and Section VI, *D*. The reliability coefficients of Table 5 are already free from the influence of Case 4, since Group *A* was not included in the sample from which the figures of this table were obtained (cf. Section VI, *C*.)

the full sample of 31, and corresponding figures for the reduced sample of 30. The differences in results for the two groups are generally fairly small, and not of a nature to alter conclusions of the study. It should be observed that the figures in Table 11 are all Pearson r's; the effect of an exceptional case upon the corresponding Bernstein r's is generally smaller (cf. Section V, *D*).

REFERENCES

1. BERNSTEIN, F. Regression and correlation evaluated by a method of partial sums. *Ann. Math. Statist.*, 1937, 8, 77-89.
2. HULL, C. L. Aptitude Testing. Yonkers, N. Y.: World Book, 1928. Pp. 535.
3. KELLEY, T. L. Statistical Method. New York: Macmillan, 1924. Pp. 390.
4. READ, KATHERINE H. Significant characteristics of preschool children as located in the Conrad Inventory. *Genet. Psychol. Monog.*, 1940, 22, 455-487.

$7.00 per volume — QUARTERLY — May, 1942
Single numbers $4.00 — Two volumes per year — Volume 25, Second Ha

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY
CARL MURCHISON

Provincetown, Massachusetts

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under the Act of March 3, 1879

Reentered as second-class matter April 15, 1937, at the post-office at Provincetown, Mass., under the Act of March 3, 1879

The Journal of Psychology

EDITED BY

Carl Murchison

The Journal of Psychology was founded at the request of many psychologists throughout the world for the purpose of establishing a medium in which publications shall be immediate and in which other publication experiments may be carried out.

(1). The subscription price will be $7 per volume.

(2). The format of each volume will be such that each subscriber may tear down the assembled volume into its component articles.

(3). Except through the fault of the author, the lag in publication should not be more than three or four weeks.

(4). The author will receive 200 reprints of his paper, and the subscriber will receive the accumulated publications bound into half volume units. The second half of each volume will contain the title page and table of contents for that volume. Each half volume unit will be mailed to its subscribers at its stated date.

(5). Each author will be charged the actual manufacturing costs of any zinc or half-tone cuts.

(6). The publication charges will be as follows:

4 page signature	....	$12
8 " "	....	24
12 " "	....	36
16 " "	....	48
20 " "	....	60
24 " "	....	72
28 " "	....	84
32 " "	....	96
36 " "	....	108
40 " "	....	120

These charges are tentative and will be adjusted as justified by the subscription list (especially by the effect of the war in China and in Europe).

(7). The accumulated units will be mailed to subscribers quarterly: January, April, July, and October. The units may vary in size, to accommodate publication needs.

Two volumes annually. Per annum $14.00; per volume $7.00; per half volume unit $4.00. Complete sets at $7.00 per volume plus transportation.

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

$7.00 per volume
Single numbers $4.00

QUARTERLY
Two volumes per year

May, 1942
Volume 25, Second Half

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special

edition, sold under the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication.

MAY, 1942

(Manuscripts are printed in the order of final acceptance)

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under the act of March 3, 1879
Reentered as second-class matter April 15, 1937, at the post-office at Provincetown, Mass., under Act of March 3, 1879.

Published as a separate and in *Genetic Psychology Monographs*, 1942, **25**, 219-287.

ARE THERE ANY INNATE BEHAVIOR TENDENCIES?*

JOHN B. SCHOOLLAND

Department of Psychology, University of Colorado, Boulder, Colorado

*Received in the Editorial Office on August 25, 1941.

ACKNOWLEDGMENT

The writer wishes to express his appreciation to Dr. K. F. Muenzinger of Colorado University for numerous suggestions and criticisms throughout all these experiments.

The writer also wishes to express his appreciation to Dr. T. H. Howells of the University of Colorado for suggesting both the apparatus and the technique employed in Experiment II of Section IV.

JOHN B. SCHOOLLAND

Department of Psychology
University of Colorado
Boulder, Colorado

I. INTRODUCTION AND STATEMENT OF PROBLEM

The experimental work hereinafter to be described was inspired by the writings of the late Prof. William McDougall of Duke University. Though considerable criticism has been levelled from time to time against his theory of innate constitution there has been little if any attempt to test his theory experimentally. With his particular list of instinctive tendencies we are here not at all concerned. Nor should these experiments be interpreted as an attempt to test or defend his particular theory of instinct in all of its ramifications. However, one of the most obvious facts of casual observation is that of differentiated behavior, in both men and animals. In spite of the profound modification of behavior through environmental influences one wonders in how far such differentiation may be rooted in innate constitution. Our concern is therefore primarily with (*a*) an experimental investigation of the existence of innately differentiated behavior tendencies, and (*b*) a suitable criterion for differentiating possible factors of innate constitution.

The problem of innate constitution is one of the greatest theoretical and practical importance. Theoretical, in that it involves an analysis and understanding of the dynamics of human and animal behavior; practical, in that our conception of the nature of these motivating forces underlies all education, political, economic, and other social philosophy. Educational theory and practice, for example, will vary widely as behavior is viewed as a product of environmental conditioning or as a discipline of dispositions rooted in innate constitution. The importance of the problem is further accentuated when we reflect that an expressed or implied solution of the problem is well-nigh inescapable. The common insistence of educators upon the problem of "individual differences," and of progressive education upon "inner motivation" afford excellent examples.

There has been a tendency to neglect research in the general field of heredity and environment as futile, confusing, and unnecessary. Since behavior must of necessity find its initial expression in a given environment it is often impossible to determine, in a given situation, just which factors have a basis in innate constitution and which in environmental conditions. Shepard and Breed's famous experiments with the pecking of chicks, for example, may be quoted in support of either point of view. Further, the developmental or organismic point of view tends to regard heredity and environment as aspects of total organization. Thus the problem of innate consti-

tution, in so far as it implies an antithetical environment, easily becomes *absorbed* in the view of organic totality. For these reasons especially, the problem of heredity and environment tends to be pigeon-holed as one of the fruitless speculations of historical psychology, and one so intrepid as to attempt to revive it is immediately subject to suspicious scrutiny.

Granting this developmental process as a function of the total organism, however, this inseparability need not become so absolute as to preclude all possibility of fruitful analysis through artificial abstraction. Concentration upon organic totality affords an intriguing and valuable view of the organism and its development in its dynamic aspects. As such, however, it can afford but little insight into the *nature* of its organization. What is it that underlies and gives direction to this "dynamic self-organization" of the individual organism? After all, it is surely legitimate to consider the "individual" in abstraction from the cosmic whole.

The psycho-physical aspects of organization (not necessarily a metaphysical differentiation) appear to be integrally related, i.e., structure, function, predisposition and perceptual ability are aspects of total organization, and as such can scarcely be differentiated as "innate" or "acquired." E.g., that there are *structural* differences between various species of animals goes without saying. Anatomical and physiological differences have always constituted a basis of major importance in the classification of animals. Further, the structural and functional aspects of animal organization are so closely associated as to equip, individually and collectively, for more or less specific types of behavior. The web feet of the duck, for example, the oily skin, the buoyant body and the flat keel, adapt him admirably for a semi-aquatic existence as compared with the structuro-functional organization of the chick. The claw-like feet of the chick, on the other hand, make it adaptable for scratching, roosting, etc., types of behavior for which the duck is not at all adapted.

Integrated with this *structuro-functional* aspect of animal organization, may we also assume at least a constitutional basis for *psychological* organization, involving a *predisposition* to characteristic individual and social types of behavior, as further aspects of total organization? That is, is there an innate tendency to respond more readily to environmental stimuli which are by nature suited to innately given structure and function? Further, is there a constitutional basis for social (gregarious) behavior, a tendency to

"associate with like kind," to respond more readily to visual and auditory stimuli from like kind rather than from unlike kind? And finally, does this gregarious tendency necessarily involve the inference of a *perceptual* ability, which functions in distinguishing like from unlike kind and is somehow integrated with the disposition to go to like kind?

Such a concept, of course, involves some kind of maturation hypothesis, a progressive developmental process of organization. For example, while it seems highly improbable that a freshly laid, fertile egg should respond to sensori stimuli (except temperature), Kuo has demonstrated, in his fascinating experiments with embryonic chicks, that they may respond to touch, pressure, sound, and light after a period of incubation. It does not seem probable that they would differentiate sounds emanating from animals of like or unlike species, however, at this period. Such differentiated response has not yet developed, is not called for by the "need" of the organism until at least some hours after hatching. If it could be demonstrated, however, that chicks and ducks, for example, *do* differ in their response to like and unlike kind shortly after hatching, and if at the same time it could be demonstrated that such differentiation was not due to environmental factors, it might be reasonably inferred that such differentiation or cognitive ability has an innate basis, and may be regarded as the "maturation" of such cognitive ability as the organism has "need," i.e., as the nature of the organism and its environment develop into a progressively integrated and enlarged unitary whole.

After hatching, however, we note a strong tendency of chicks and ducks to associate with like kind. Though this is often attributed to environmental circumstances it is quite possible that such tendencies and perceptual abilities constitute an integral aspect of organization viewed as a whole. Thus Schjelderup-Ebbe says that "each bird *learns more easily to distinguish* individuals of its own species than it does to distinguish individuals of another species."

In a thought-provoking discussion of what he terms the "differential point of view," Howells points out what promises to be a most effective method of coping with the elusive though persistent problems of heredity and environment. He says that the chief obstacle in the way of dealing with the problem lies in the conventional concept of heredity as "static and absolute" rather than "dynamic and comparative." Actually there is no way of determining when the "outer" factors begin to interact with the "inner" factors.

As the problems of heredity are comparative problems throughout, all practical definitions must be relative or differential only. What is required, therefore, is some arbitrary *point of differentiation* and some *method of differentiating* the factors which then may be designated as representing the "inner" and the "outer" factors. As the pre-natal conditions are largely beyond observation and it is easier to control the outer factors from birth on, he selects "birth" as the obvious *point* of differentiation. Heredity may then be defined as the "dynamic pattern of differential factors which are present in the organism at birth." He suggests that the "dynamic" criterion of "ease of learning" be substituted for the older equivocal and "static" criterion of "appearance without practice" or "unlearnedness" as the *method* of differentiating innate factors. In our discussions we shall adopt the general import of the term "ease of learning." As our experiments are not concerned with typical "learning" situations, however (except for our last experiment), we prefer a modification of this criterion which for our purpose might better be termed "facility of response" in an equivalent situation.

From this point of view there can scarcely be any objection to setting up an arbitrary point in the developmental process and to determine to what degree the development up to this point exerts a determining influence upon subsequent development. Those factors present in the organism at birth, which represent development *up to this point* and which are at the same time partial determinants of subsequent behavior we shall regard as *innate*, and as *representing innate constitution*. Those conditions surrounding the organism at birth, and which are partial determinants of *further* development we shall regard as *external*, and as *representing environmental influence*.

Given this point of departure—the organism at birth—we now proceed to a discussion of the variables involved in an experimental investigation of the factors of heredity and environment.

A. Two Sets of Variables—Heredity and Environment

In the manipulation of any two sets of variables, two, and only two, methods are possible. Thus in our consideration of the two variables, heredity and environment, we have two possible methods of *experimental* approach, each with its own peculiar advantages and disadvantages.[1]

---

[1]Psychological literature has contributed much in the nature of *uncontrolled* observations of animal behavior, with which we are here not particularly concerned.

1. *Heredity May Be Held Constant While Environment Is Varied*

One of the best known experiments of this type is Kuo's oft-quoted experiment of the "Genesis of the Cat's response to the Rat." Kuo varied the diet and other environmental conditions of the cats and then noted their behavior with respect to rats. The results indicated clearly that environment played an important rôle with respect to cat-rat behavior. The experiment is particularly interesting in that it has so often been quoted, *erroneously*, as evidence for the non-existence of innate tendencies. This does not mean, of course, that there *is* such a thing as a cat-rat killing tendency. It means only that where the heredity factor is held constant one is warranted in drawing conclusions *only* with respect to the influence of varied environmental factors, not with respect to heredity factors. An important exception must be noted, however, in the experimental work being carried on with identical twins. In these researches the environmental factor is a variable, but, instead of being designed to determine how environment may *change* native constitution, they are carried on with the purpose of determining how *constant* the factor of innate constitution may *remain* under greatly different environmental conditions. The heredity factor being assumed in this instance, the aim is not so much to determine its nature as it is to determine its strength.

2. *Environment May Be Held Constant While Heredity Is Varied*

This method, that of keeping the environment factor constant while varying the heredity factor, is essentially the one proposed by Howells, and is the one employed in the present experiments. Barring the exception just referred to, it is the only method which can afford direct, compelling evidence bearing on innate constitutional factors. For, by maintaining the environment constant, from birth, any differentiation in behavior must be regarded as rooted in innate constitutional differences.

While we are primarily concerned with the nature of human heredity, and hence with its bearing on human individual and social (and educational) psychology, it is often difficult to experiment with human material. This difficulty led to the selection of chicks and ducks as sufficiently differentiated animals, and which could be easily manipulated, and whose environment could be readily controlled and held constant from birth. It is confidently assumed, however, that if the validity of the concept of psychological inheri-

tance can be established on a lower phylogenetic level it can be applied equally to human inheritance and human nature, even though it be obscured by the longer period of "infancy." Such a concept should be of the utmost importance for the nature of human individual and social organization and educational procedure.

B. Statement of Problem

If chicks and ducks, hatched at the same time, in the same incubator, in equal numbers of each, and reared in the same environment from birth, and thus with "equal opportunity to learn," or better, "with equal opportunity for differential response," are able to distinguish like from unlike kind, and tend more readily to associate with like rather than with unlike kind, such ability and tendency, appearing in the face of equivalent environment, cannot be interpreted in terms of environmental factors only but must be regarded as aspects of total "chick" and "duck" organization, and hence as having their basis in innate constitution.

C. Plan of the Study

The experimental section of this discussion is divided into four parts. (*a*) "Field Studies" is an account of preliminary controlled observations. These were largely of an exploratory nature and were conducted as a basis for more elaborate experimentation. (*b*) "The Tendency of Newly-Hatched Chicks and Ducks to Approach Like Kind," deals with the spontaneous social behavior of animals in a "free choice" situation. The writer regards this experiment as the most important of this series. (*c*) "Affective Response to Simultaneous Shocking of Like and Unlike Animals under Equivalent Environmental Conditions," deals with the initial social responses of animals to cries of like and unlike kind. (*d*) The fourth part deals with the "Positive and Negative Responses of Chicks and Ducks to Their Initial Perception of Water."

II. FIELD STUDIES

A. Period of Incubation

First, of course, was the necessity of determining the period required for incubation, so as to insure as nearly simultaneous a hatch of chicks and ducks as possible. This period seemed to differ slightly with breeds and perhaps with seasons.

B. Initial Response to Drinking Water

Each chick and duck upon being removed from the brooder for the first experience with food and water was placed individually in a small wire compartment, with a small round tin, three inches across and ¼ in. deep, filled with water and placed in the center of the compartment. A brief characteristic description for each species will suffice to illustrate the different behavior elicited upon initial perception of water to drink.

> *Duck No. 3 (a sturdy individual—a drake).* Walks about, peeping lustily. At 4 sec. steps in water. Noticed and sought drops on floor at edge of tin. At ¼ min. stepped in tin, stood in and sought to drink from the floor; 5 sec. later found water and drank 7 times, continued to seek water on floor near tin (to drink) 7 times—drank twice from tin, then twice from floor, drank 4 times—*went through motions of drinking* water from the tin *four times, without actually* touching water—drank twice from tin—went through motions *twice* without touching water—drank from tin twice—went through motions once. After 11 min. stood in water while making drinking motions *four times,* then made the motions *nine times* touching wet floor each time, but no water. Then drank 5 times while still standing in the water, then *four times* making the motions without touching anything. Then 26 times in succession, touching floor (now wet in places). Then 11 times on the floor, then 3 actual drinks—then began peeping again. After 16 min. (from start) "drank" 7 times from the floor.
>
> One min. later put bill in water twice, now making *"sucking noises"* and shaking the head. Then 5 times made similar "noises" in the water, once on floor. Then took a drink. Began walking around, for 2 min., without further drinking. Then walked past water once, then through, without paying attention to it. Then walks through water 4 times, giving no apparent attention to it. Peeps and makes "searching" noises in water on floor again (though care has been taken to keep floor

> and water free from all particles of food, etc.) After 21 min. is taken out, since continues to walk about, peeping.
>
> Apparently the perception of the water called forth both the "drinking responses" and the characteristic "straining movements" with the bill, upon its very first presentation.
>
> Contrasted with this type of response is that of the chick, No. 2. After ½ min. stepped in water with one foot and stood thus for 10 sec. Gave no attention whatever to water. Stepped in three times in 1 and ½ min. Pecked once at the surface of the water, did not drink. After 4 min., stepped in again and fell in it. Took one drink. One min. later stepped in again, giving no attention. Removed after 8 min. (from start), having given no more attention to the water.

Chicks and ducks differ greatly in their initial response to water. Difference in sensitivity of the feet doubtless plays an important part, though it must be regarded as but an aspect of the total integrated "nature" of the duck.

C. Initial Response to Food

Food pan (containing dry mash), and water tin (shallow tin cover, as above) were placed 2½ ft. apart. *Chicks* gather about food pan particularly, eating constantly, and drinking only occasionally.

Ducks gather constantly about the water tin. Then they go to the food pan, peck at the food four or five times (sometimes only once or twice) and then *run* rapidly to the water, drink, and *run* back to the food pan, repeating this with considerable regularity for as long as five or six minutes without perceptible let-up. This was very characteristic behavior of the ducks with any kind of food, though slightly more so with the dry mash. The impulsive, "hurried" behavior of the ducks is strongly in evidence, but cannot alone account for behavioral differences. Observations were made at initial feeding, when 36 hours old.

D. Drinking Responses to Other Presentations

A mirror (8 x 12 in.) was placed on the lawn. The ducks approached it almost immediately, pushing their bills across it repeatedly as they walked, and sometimes along its edge. To all appearances, their behavior was exactly similar to that in shallow water. A circular piece of white oil-cloth (3 ft. in diam.) placed on the

lawn called forth similar responses from the ducks. The chicks, however, in both instances either ignored or pecked only casually at the surface (age, four weeks).

E. Flight in Response to Cries of Distress

On one occasion seven chicks and seven ducks (five days old) were feeding on chopped lettuce, of which the ducks especially are very fond. They were all on a narrow platform, 1 x 3 ft., in front of the brooder. The observer quite coincidentally picked up a chick from another group. Accidentally it slipped through his fingers and was caught by one foot. It cried out lustily. Immediately the seven chicks ran into the brooder, while the seven ducks continued eating.

F. Food Preferences

Ducks are very fond of chopped lettuce from the very first day. They eat it eagerly, often in preference to other foods. Chicks on the other hand, are very found of the dry "starter mash." The ducks care less for this, and when eating it they wash it down with frequent drinks of water.

The ducks seek particles of food at the bottom of their water tin. The chicks content themselves with pecking at particles floating on the surface. These differences in *initial* behavior may again be intimately correlated with structural differences, with the difference between bill and beak, and perhaps also differences in vision. Incidentally, the ducks are far more impulsive and energetic in their eating habits, as they are in most phases of their behavior.

G. Preference for Duck or Hen Egg as Food

The difference in the quality of pre-natal nutrition suggested a possible preference for egg, especially of like species, as food. Accordingly, unfertilized duck and hen eggs were prepared for their *first* feeding, when the animals were 36 hours old. The eggs were boiled for one hour, cut cross-ways, weighed, and placed in circular holes cut in a small board 3 x 8 x 1 in. (see Figure 1). These boards were then placed in separate compartments containing three ducks and three chicks respectively. Here they were left for 5 min. periods, the position of the egg in each board being reversed every 30 seconds by turning the board through an angle of 180 degrees.

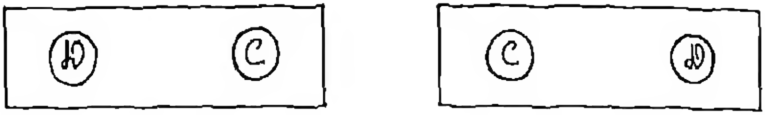

FIGURE 1

TWO BOARDS CONTAINING COMPLEMENTARY HALVES OF DUCK EGG (*D*) AND HEN EGG (*C*) FOR DETERMINING EGG PREFERENCE

The difference in weight was then noted, with the following results:

1. The preference, in each instance, both for chicks and ducks, was for the egg yolk. The white was eaten only after the yolk had disappeared.
2. Little preference was shown for either kind of egg.
3. Chicks often ignored the egg throughout the whole five-minute period. The ducks invariably attacked the egg and ate it greedily.
4. The ducks (both at this age and later), seemed to prefer the egg to other kinds of food. This was not nearly so pronounced with the chicks.
5. The chicks, when they did eat, usually gathered about the same egg, pecked deliberately, and seldom left to drink from the water fountain placed at a distance of about 2½ ft. from the egg. The ducks on the other hand, ate greedily, *running* for a drink of water with every three or four pecks at food, returning haphazardly to one or other of the eggs. This, together with their impulsiveness and greediness, made the technique rather difficult. Occasionally they would transfer from one egg to another, but usually only when a drink intervened.
6. Chicks which at first ignored the egg would soon approach when placed with ducks already eating. They engaged preferably, however, in pecking at small particles lying about than at the egg itself, suggesting perhaps visual and other differences. In 90 min. 84 ducks ate 89 gr. duck egg and 79 gr. hen egg. In 45 min. 39 chicks ate 21 gr. duck egg and 20 gr. hen egg.

For want of time and difficulty of technique this experiment was not carried through to completion. But it offers interesting possibilities. It is possible, too, that the boiling of the eggs was an error.

H. Initial Response to Water Tank at One Week

At one week of age these same chicks and ducks (reared together from birth) were placed, two each at one time, in a cage, 2 x 4 ft., with a water bath, 11 x 17 in., and 5 in. deep, near one end of the cage. None of the animals had ever experienced water other than that in their inverted drinking fountain, which was just enough to wet the bill.

Invariably the ducks hastened to the water, launching forth without apparent hesitation, and swam about in apparent fearlessness. They usually emerged from the first plunge within 15 seconds, only to return within as many more. Several even dived completely and repeatedly under the surface of the water, emerging a foot or so away and throwing water with their heads.

The chicks would sometimes run to the water *with the ducks,* sometimes approach more casually, then walk near the edge, and withdraw. None made any attempt to enter. One, however, which inadvertently slipped and fell in, struggled desperately to escape—conduct very different from that of the ducks entering the water.

It should be noted in this connection, however, that a considerable number of ducks, some of one week and some of six weeks of age, when *placed* in water for the first time (never having been in water before), invariably *walked,* if the shallowness of the water permitted. It was only as the depth of the water was gradually increased that they gradully "floated," rising higher on their toes, until, when the toes could no longer reach, they pulled the legs up in swimming position and swam about.

Although the initial response of chicks and ducks to the perception of water differs greatly, even when presented in groups of equal numbers of each, it may be, of course, that the natural greater buoyancy of the duck plays a considerable rôle in further behavior, once the animal is launched. This serves once more to emphasize the integration of anatomical characteristics.

I. Initial Response to Water at Five Months

Three ducks and three chickens, *hatched together and reared together, from birth, in close confinement,* had never experienced more water at one time than could be contained in a can five inches across and three and one-half inches deep. (These animals had not been used in any previous experiments.) James' famous theory of "transi-

toriness of instincts" suggested an experiment. On a warm midsummer afternoon the animals were given fresh, cool water to drink. When they had quenched their thirst, an empty galvanized-iron tub was placed upon a box one foot high in their runway. Then it was filled very quietly from a garden hose. Next it was set upon the ground, empty, and then filled as before. Later it was placed in the ground, so that the level of the water was even with the ground, but with wire poultry netting entirely encircling it. The behavior of both chickens and ducks was noted after each presentation. All were somewhat "curious" at first. When the tub was placed upon the ground, filled, the ducks began to show excitement as they walked about it, in strong contrast to the idly "curious" behavior of the chickens. When the tub was placed *in* the ground, but with wire preventing entry, the ducks became greatly excited. They made "searching" movements with their bills along its edge. While the chickens walked indifferently at a distance, the former disported themselves, "ducking" their heads and bodies up and down, rising high on their legs, and flapping their wings, and generally acting much like ducks enjoying a long-delayed swim. After some 10 minutes of this they walked off, and spent some minutes *preening their feathers* from neck to tail, just as though they had had a swim, *though the feathers were perfectly dry!* The difference in behavior is all the more significant in that the ducks and chickens had been reared *in the same environment from birth,* with *equal opportunity* to learn, and yet showed this differentiated behavior. This behavior was recorded throughout on motion-picture film.

Another interesting type of behavior was differentiated under these equivalent environmental conditions. The ducks walked excitedly around the tub, bobbing their heads up and down. At the height of their excitement, apparently, two of the ducks engaged in the mating act, observed in this group for the first time. This "bobbing" of the head at the perception of a pond of water is very similar to the behavior which customarily precedes or accompanies the mating act. This same "bobbing" was also observed among younger ducks, from six weeks of age onward, when they were thirsty and heard running water. Thus one type of physical expression seems to be closely associated with three major activities, those of swimming, mating, and drinking.

The wire screen is now removed. Ducks excitedly drink and put their heads far beneath the water (their first opportunity).

After 30 seconds one duck "drops" in the water and swims across, emerging on the other side in about five seconds. After $2\frac{1}{2}$ minutes another enters, swims across, and emerges likewise in about five seconds. Three minutes later the drake enters and emerges similarly. Egress is accompanied by rising high on the legs and vigorously flapping the wings. They continue flapping and preening for 15 minutes more, during which time three other entries are made. The chickens meanwhile are several feet from the water, giving little if any heed.

Food is then thrown into the water (cracked grain). The ducks eagerly "screen" for it on the surface. The hens approach and peck at floating particles. When the surface particles have mostly disappeared the chickens walk away. The ducks continue to seek food by putting their heads far under water for considerable periods. Soon the drake enters the water and continues this method of feeding. The water is removed after 20 minutes.

J. Response to Third Dimensional Relations

The spontaneous launching forth of the ducks into the water suggests another related type of behavior. How would chicks and ducks behave in descending from a slight elevation?

Thirteen chicks and 13 ducks (hatched and reared simultaneously, in equal numbers of each) were placed one at a time on a box 2 in. in height, immediately after being taken from the incubator, 12 to 24 hours after hatching. It was their first exit from the incubator, never yet having experienced either food or water. The box and the surrounding area beneath were completely covered with white paper, to avoid brightness difference. Two animals of each species were placed in a small cage at a distance of two feet to attract the individual on the box in a given direction.

The ducks immediately ran off the box as though running on a level surface. The chicks quite similarly, but alighting with greater success.

Next, a box 6 in. high was substituted for the other. Again the ducks launched forth, with very little if any hesitation. The chicks, on the other hand, paused at the edge, frequently walking for a short distance along the edge, then deliberately *jumping* down.

Finally a box 1 ft. high was substituted, covered as before. The ducks gave evidence of hesitation in some cases, but usually plunged impulsively over the edge. The chicks showed even more de-

liberation than before, pausing as long as one-half minute, and walking back and forth along the edge before *jumping* down.

This experiment again indicates a difference in visual quality, or a native impulsiveness, or both. And it may be related to the impulsive launching forth upon the water as in the foregoing experiment. In any event, it again appears to give evidence of an organization of factors into an integral whole. Also, under the conditions imposed (of environmental equivalence, from birth) it affords evidence for innate differentiation of behavior.

K. Competition for Food, and Cessation of Flight Upon Swallowing

Another observation merits a brief note. The greater appetite and greediness on the part of the ducks for food and water has already been noted. Coinciding with this is the tendency of the duck to gulp his food quickly when another attempts to take it from him. The chick tends much more readily to run with it, frequently giving more attention to possible pursuit than to the swallowing of the food. This suggests another inter-relation of anatomical and psychological factors, as the duck is less adequately equipped for pecking his food to pieces and is obviously handicapped in locomotion, especially on land.

Another interesting phenomenon is the sudden cessation of escape activity, especially on the part of the duck, as soon as the morsel competed for has been *swallowed*. A precipitate flight may be terminated. Doubtless this is learned, but very readily.

L. Different Responses to Water and Dust Presented Together

Another experiment provided an environment with two varied aspects, in order to differentiate behavior of chicks and ducks reared in an equivalent environment from birth. At two weeks of age, a group of animals (not used in any previous experiment), four chicks and four ducks, were placed in a wire enclosure in a shaded spot, on a hot afternoon. A circular strip of wire netting enclosed a dry, hot, dusty piece of ground, 2½ feet in diameter. A small place about eight inches in diameter had been scooped out at one side and filled with water. None of the animals had been on bare soil before, but were closely confined either on a board floor or on a thick grassy plot.

A description of one group will suffice for the several used. The ducks immediately approached the water, entered and engaged in pushing their bills through the water and mud and splashing generally. Almost without exception the ducks remained in or near the mud puddle throughout the 10 minutes of observation.

Of the chicks, however, not one entered the mud, though several approached. Within one minute two were dusting themselves vigorously in characteristic fashion, enjoying the experience for the first time. During the next minute another joined them. Thus they busied themselves throughout the whole period, at times stretching out neck and leg on one side, closing the eyes, basking in the sunshine then, after some more dusting activity, stretching out on the other side. (The ducks did not attempt to dust themselves.) The fourth chick walked about between, mingling first with the one and then with the other, occasionally stepping in the mud, but neither lying in it (as the ducks were doing about him), nor lying in the dust.

M. Sex Behavior of Adults—Reared in Equivalent Environment

In the case of the three ducks and three chickens reared together in close confinement from birth for six months, the ducks occasionally showed some attention to a hen, treading a hen at times. Considerable more attention, however, was given to the female of his own species. A significant difference lay in the fact that the hen was caught and forcibly held by the neck feathers or comb, while in the preliminary courtship of the duck, preceding the act of mating, there seems to be a very decided mutual understanding and cooperation.

It is not clear that the drake differentiated between the sexes of the chickens. He seemed to behave quite similarly toward each. The response of the hen or cockerel itself seemed to determine his behavior, the former being more submissive, the latter more vigorously resisting his approach or advances. It was observed, however, that the drake would sometimes persecute the cockerels, driving them about furiously and pulling out feathers, often passing the hen without molestation in doing so. This need not necessarily be a cognition of sex differences; it may be explained, perhaps, in the previous relation between the drake and cockerels. The cockerels on the other hand, did not attempt to molest the female ducks—

perhaps due to the aggressiveness of the drake. Behavior differentiated through equivalent environmental experience would seem to indicate, however, an innate basis for greater submissiveness on the part of the female and greater aggressiveness on the part of the male.

N. Sex Behavior of Adults—in a Non-Equivalent Environment

In another group of fowls, in which no attempt was made to maintain equivalent environmental conditions, there were eight cockerels, eight pullets, two ducks, and two drakes, all but the drakes being of equal age, however (seven months). The drakes were two months older. In this situation, perhaps because of the comparatively large number of cockerels, much attention was given by the latter to the ducks. Here again, it was not clear that there was real cognitive differentiation of the sexes, and the act of copulation seems to have been incomplete, due to the difference in length of body. The cockerels frequently sidled up to the drakes as well as to the ducks. As in the foregoing observations, the subsequent behavior seemed to be determined not only by the aggressor but by the sex of the animal approached. The ducks usually attempted to escape, but were pursued and caught. The drakes, on the other hand, showed fight and were let alone, sometimes defending a duck as well. In one instance a drake caught a cockerel by the wing, whirled him rapidly in a circle four or five times, and then let him go, whereupon the slightly dazed cockerel staggered in the direction of release and desisted.

The hopeless outnumbering of the ducks by the cockerels seems to be largely responsible for the phenomena. Several large poultrymen informed us that they have frequently observed drakes to tread hens, but had never heard of cockerels or roosters treading ducks. The ducks, who characteristically engage in a period of courtship with the drakes before mating (a differentiated type of behavior doubtless associated with structural and cognitive differences), often resisted the cockerel. This might have been successful were it not for the prompt arrival of overwhelming reinforcements of other cockerels. Helpless and apparently hopeless before this onslaught, the duck would attempt flight and be promptly overtaken. The cockerels frequently attempted to hold the duck by the nostrils.

The drakes would sometimes interfere, with some degree of success, though they too, were hopelessly outnumbered.

As the number of the cockerels was gradually reduced, however, it was observed that they molested the ducks less and less. When their number was reduced to two or three they seldom disturbed the ducks. And when one, the most aggressive cockerel of them all, was left alone with the eight pullets and four ducks, he was not observed in three months to have attempted to tread one of the ducks—even during the last month, when the drakes had been removed also.

Cinematographic records were made of a number of these observations, for example Nos. 2, 3, 4, 7, 9, 10, 12 and 14. Similar records were also made of other features of subsequent experimentation.

In all these experiments the differentiated behavior of chicks and ducks appeared in the face of equivalent environmental experience from birth and with "equal opportunity to learn." It must be concluded, therefore, that the differentiated behavior, called forth upon initial perception, cannot be accounted for adequately in terms of environmental experience but must be regarded as rooted in innate constitution.

III. EXPERIMENTAL STUDIES: SOCIAL BEHAVIOR IN CHOICE SITUATION

Much of the chick-duck behavior in our earlier observations was closely integrated with structuro-functional differences. Though we regard these differences as aspects of "total chick-duck organization" it seemed desirable to eliminate these factors as much as possible in favor of the more purely psychological. It was therefore decided to employ a type of experimentation involving perceptual and discriminatory abilities on the one hand and social (gregarious) tendencies on the other. These abilities and tendencies appeared to be admirably integrated in the "tendency to associate with like kind." While these abilities cannot, of course, be directly observed, they may be readily inferred from behavior.

A. Experiment I: Tendency of Newly-Hatched Chicks and Ducks to Approach Like Kind, Gregarious Tendency

1. *Problem*

How do chicks and ducks, hatched and reared in an equivalent environment, hence with "equal opportunity to learn," respond to the presence of like and unlike animals during the first 10 days after hatching?

2. *Materials*

Thirty chicks (White Leghorns) and 30 ducks (White Pekins) were used in this experiment, each banded and numbered. After some experimentation to determine the time required for these breeds, incubation was begun in separate incubators in order to maintain adequate humidity for each, incubation of duck eggs being begun one week earlier than that of the hen eggs. Two days before hatching all the eggs were placed in the same incubator. During the hatching period care was taken to keep the number of hatched of each kind as nearly constant as possible. After hatching, chicks and ducks were reared in close confinement in equal numbers of each, in order to insure equal contact (visual, auditory, tactual, olfactory, etc.) between like and unlike kind.

In order to provide as "natural" a setting for the experiment as possible, with a maximum of freedom from artificial stimulation, the following device was employed. A "Choice Cage" (Figure 2) was constructed on a wooden base four feet long by two feet wide, enclosed on all sides by one-half inch square wire mesh, one foot high.

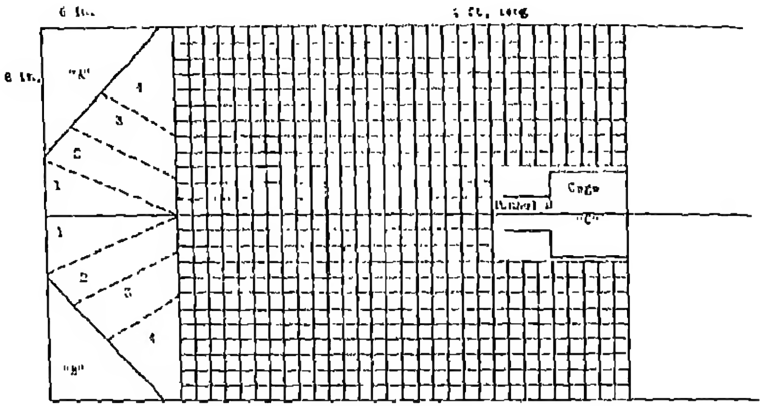

FIGURE 2

Diagram of "Choice Cage"

Experiment I: Cognition of kind inferred from choice in going to a single chick or duck.

Cages "*A*" and "*B*" hold one animal each, a chick and duck respectively, alternating with each series.

Cage "*C*," from which single animals are released, and observed for their behavior in going to like or unlike kind, or to neither.

Animals are released from "*C*" through sliding door "*D*" and tunnel.

Spaces numbered 1 to 4 indicate area of approach.

Floor of cage is marked off in one-inch squares to facilitate measurement of degree of movement, if desired.

At one end the two corners of this cage were enclosed by wire mesh, making triangular compartments, "*A*" and "*B*," each eight by eight by twelve inches. In the middle of the opposite side there was another wire compartment, "*C*," six inches square and six inches high, made of the same type wire mesh. In front of this compartment was an attached wire tunnel of similar material, three inches wide, so that when animals were released from "*C*" they were obliged to make their exit in a direction midway between "*A*" and "*B*." Between tunnel and compartment "*C*" was a section of wire screen (same material) which could be raised horizontally to release the animal confined in "*C*." The sides, back and top of "*C*" was enclosed in a hood of heavy black construction paper. The floor of the entire cage was enameled white and marked off in inch squares with India ink to facilitate easy measurement in any direction.

3. *Procedure*

A chick and duck were placed in compartments *"A"* and *"B"* respectively. A third animal was placed in compartment *"C."* After an orientation period of 30 seconds the animal in *"C"* was released by raising the slide. Its behavior with respect to the animals in *"A"* and *"B"* was then noted.

Ducks and chicks were placed alternately in *"C"* until each had had his turn in *"C."* Cue animals in *"A"* and *"B"* were changed from time to time in order to equate somewhat the "stimulus value" of the cue animals.

Two such series were conducted daily, in the second daily series the kind of animal in *"A"* and *"B"* being reversed. The double series were conducted daily, morning and afternoon, beginning when the animals were one day old and continuing for 10 days.

Behavior was noted with respect to their approach to *"A"* or *"B"* or their failure to approach either. If the released animal approached to within two inches of the white space containing cage *"A"* it was regarded as having approached the animal in *"A."* Similarly for the animal in *"B."* If it failed to approach either within thirty seconds it was regarded as approaching neither.

Sixty trials were made for ducks and chicks respectively (twice for each animal)—a total of 600 trials for each species for the 10 consecutive days following hatching.

4. *Results*

The results for a total of 600 trials in the 10 daily series were as follows:

Approaching Like or Unlike Animal

Duck approaching duck:	392 times or 65.2% of trials.
Duck approaching chick:	164 times or 27.5% of trials.
Duck approaching neither:	44 times or 7.3% of trials.
Chick approaching chick:	236 times or 39.2% of trials.
Chick approaching duck:	193 times or 32.3% of trials.
Chick approaching neither:	171 times or 28.5% of trials.

The above figures may be recombined to indicate the relative *strength* of gregarious tendency of chicks and ducks in another way, as in the tendency to "approach another animal" whether of like or unlike kind, as follows:

Duck approaching another animal:	556 times or 92.3% of trials.
Duck approaching neither animal:	44 times or 7.7% of trials.

Chick approaching another animal: 428 times or 71.3% of trials.
Chick approaching neither animal: 172 times or 28.7% of trials.

The behavior of ducks and chicks for *each of the 10 daily series* is shown in Table 1.

TABLE 1

SHOWING BEHAVIOR OF DUCKS AND CHICKS FOR EACH OF THE TEN DAILY SERIES

Day	Duck's behavior Direction	Times	%	Chick's behavior Direction	Times	%
			A. In approaching a duck or chick			
1.	D to D	35	58.3%	C to C	13	21.7%
	D to C	19	31.7%	C to D	21	35.0%
	D to —	6	10.0%	C to —	26	43.3%
2.	D to D	40	66.6%	C to C	24	40.0%
	D to C	13	21.6%	C to D	15	25.0%
	D to —	7	11.8%	C to —	21	35.0%
3.	D to D	35	58.3%	C to C	21	35.0%
	D to C	19	31.7%	C to D	18	30.0%
	D to —	6	10.0%	C to —	21	35.0%
4.	D to D	37	61.7%	C to C	25	41.7%
	D to C	17	28.3%	C to D	22	36.6%
	D to —	6	10.0%	C to —	13	21.7%
5.	D to D	39	65.0%	C to C	26	43.3%
	D to C	20	33.3%	C to D	19	31.7%
	D to —	1	1.7%	C to —	15	25.0%
6.	D to D	46	76.7%	C to C	30	50.0%
	D to C	12	20.0%	C to D	18	30.0%
	D to —	2	3.3%	C to —	12	20.0%
7.	D to D	42	70.0%	C to C	27	45.0%
	D to C	17	28.3%	C to D	21	35.0%
	D to —	1	1.7%	C to —	12	20.0%
8.	D to D	38	63.3%	C to C	23	38.3%
	D to C	17	28.4%	C to D	20	33.3%
	D to —	5	8.3%	C to —	17	28.4%
9.	D to D	38	63.3%	C to C	23	38.3%
	D to C	19	31.7%	C to D	13	21.7%
	D to —	3	5.0%	C to —	24	40.0%
10.	D to D	42	70.0%	C to C	23	38.4%
	D to C	11	18.3%	C to D	26	43.3%
	D to —	7	11.7%	C to —	11	18.3%
			Totals for the ten days			
Duck to duck		392	65.3%	Chick to chick	235	39.1%
Duck to chick		164	27.3%	Chick to duck	193	32.1%
Duck to neither		44	7.4%	Chick to neither	172	28.7%

TABLE 1 (*continued*)

B. In approaching "another animal"

(Combined approach to like and unlike animal in above table.)

Ducks to "another animal" Day	Times	%	*Chicks to "another animal"* Day	Times	%
1.	54	90.0%	1.	34	56.6%
2.	53	83.3%	2.	39	65.0%
3.	54	90.0%	3.	39	65.0%
4.	54	90.0%	4.	47	78.3%
5.	59	98.3%	5.	45	75.0%
6.	58	96.6%	6.	48	80.0%
7.	59	98.3%	7.	48	80.0%
8.	55	91.7%	8.	43	71.7%
9.	57	95.0%	9.	36	60.0%
10.	53	88.3%	10.	49	81.7%
Average	55.6	92.7%		42.8	71.3%

The behavior of *each of the 30 ducks and chicks,* in each of the 10 daily series is shown in Tables 2 and 3 respectively.

The relative strength of the tendency of chicks and ducks to approach like kind is shown graphically in Figures 3 and 4; and the relative strength of tendency to approach "another animal" in Figures 5 and 6.

Both ducks and chicks, hatched and reared in an environment held constant for all from birth, display a distinct (gregarious) tendency to go to like kind rather than to unlike kind. The ducks, however, show a stronger gregarious tendency than do the chicks, both in going to like kind and in going to "another animal."

5. *Conclusions*

In view of this behavior, differentiated in the face of equivalent environmental conditions from birth, we conclude:

1. That chicks and ducks, in their very nature, exhibit a *gregarious tendency,* and that this tendency cannot be adequately interpreted in terms of differentiated environment but must be regarded *as rooted in innate constitution.*

2. That the tendency to go to like kind involves the necessary inference of an ability to discriminate or to cognize like kind, and that this psychological aspect of behavior is therefore an integral part of the total organization of these animals.

3. That, rather than a clear-cut distinction of acquired and innate behavior, the relative "ease" with which such differentiated

TABLE 2

TABLE FOR DUCKS: SHOWING THE BEHAVIOR OF EACH OF THE DUCKS IN EACH OF THE TEN DAILY SERIES OF TRIALS
Approach to Duck (*D*); Chick (*C*); or Neither (—) is shown for each of the two daily trials.

Duck No.	Day 1		2		3		4		5		6		7		8		9		10		Total to Duck	Total to Chick
	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to		
1.	C	C	C	D	C	D	C	D	C	D	D	D	D	D	—	—	C	D	C	D	10	5
2.	C	—	—	—	D	C	C	C	C	C	D	C	D	C	D	C	D	D	D	—	7	12
3.	D	C	D	D	D	C	C	D	D	D	D	D	C	D	C	C	D	D	—	D	13	8
4.	D	D	D	D	D	D	D	D	D	C	D	D	D	D	D	D	D	D	C	D	18	2
5.	D	D	D	D	D	C	—	D	C	C	—	D	C	D	C	—	C	D	D	D	11	6
6.	C	C	D	C	D	C	D	C	D	D	D	C	D	D	D	D	C	D	D	D	13	7
7.	C	D	D	D	D	D	C	—	C	C	D	D	C	D	D	D	C	D	—	—	11	6
8.	D	C	D	C	D	C	D	C	C	D	C	C	D	D	D	—	D	C	D	C	10	9
9.	D	C	D	D	D	C	D	C	D	C	D	D	D	C	D	D	D	D	D	D	15	5
10.	D	C	D	D	—	D	D	D	D	D	D	D	D	D	D	D	C	D	D	C	16	3
11.	D	D	D	D	D	—	D	—	D	C	—	D	C	D	D	D	C	C	C	D	12	5
12.	C	D	D	C	D	D	D	D	D	D	D	D	D	C	D	D	D	D	D	D	17	3
13.	D	C	C	C	C	D	C	D	C	C	C	D	C	D	D	—	D	D	C	D	9	10
14.	D	D	D	D	C	D	D	D	C	D	C	D	D	D	C	D	C	D	C	D	14	6
15.	D	C	D	C	D	—	D	D	D	C	D	D	D	C	D	C	D	C	D	D	13	6
16.	D	D	—	—	—	—	D	—	—	D	D	D	D	C	D	—	D	—	D	D	11	1
17.	D	D	D	D	D	D	D	D	C	D	C	D	C	C	D	D	C	D	D	C	14	6
18.	D	D	C	D	C	D	C	D	C	D	D	D	D	D	C	D	C	D	C	D	13	7
19.	—	D	D	D	D	—	D	—	D	D	D	D	D	D	D	C	C	D	D	D	15	3
20.	D	D	D	D	C	D	D	D	D	D	D	D	D	D	D	D	D	—	D	C	17	2
21.	—	—	C	—	C	C	C	D	C	D	C	D	D	C	D	D	D	D	D	C	9	8
22.	C	D	D	D	D	D	C	D	D	D	C	D	C	D	C	D	C	D	D	D	14	6
23.	C	D	D	D	D	D	—	D	D	D	C	D	D	D	C	D	—	D	—	D	14	5
24.	D	D	D	D	D	D	D	C	D	D	D	D	D	—	D	—	D	C	D	C	15	5
25.	—	D	C	D	D	D	C	D	D	D	D	D	D	D	D	D	D	D	D	D	17	2
26.	D	—	D	—	D	D	C	D	D	D	D	D	D	C	C	D	D	D	D	D	15	3
27.	C	D	D	D	C	C	D	D	D	C	C	D	C	D	C	D	D	C	D	D	12	8
28.	C	C	C	D	C	C	C	D	D	C	D	D	D	D	C	D	D	D	D	D	12	8
29.	D	D	C	—	C	C	D	D	D	D	D	D	C	D	C	D	D	C	D	—	12	6
30.	D	C	D	C	D	D	C	D	D	D	D	C	D	D	C	D	C	C	D	D	13	7

Totals for the Ten Days

Duck to duck (D)	392 times	65.3%
Duck to chick (C)	164 times	27.3%
Duck to neither (—)	44 times	27.4%

TABLE 3

TABLE FOR CHICKS: SHOWING THE BEHAVIOR OF EACH OF THE CHICKS IN EACH OF THE TEN DAILY SERIES OF TRIALS
Approach to Chick (*C*); Duck (*D*); or Neither (—) is shown for each of the two daily series.

	Day																					
Duck	1		2		3		4		5		6		7		8		9		10		Total to	
No.	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	Chick	Duck
1.	D	—	—	C	C	—	C	D	C	D	C	C	C	D	C	—	C	C	C	—	11	4
2.	—	—	C	C	C	C	D	C	C	C	D	C	D	C	—	D	D	—	D	—	9	6
3.	—	—	C	—	C	—	C	—	—	—	D	C	D	—	D	D	—	—	—	C	5	4
4.	—	D	C	D	D	C	D	C	D	D	C	C	D	C	—	D	C	C	D	D	8	10
5.	D	—	D	C	—	C	C	C	C	D	D	—	C	C	C	C	C	C	C	D	14	4
6.	—	C	D	C	C	—	D	C	D	C	C	C	C	C	C	C	—	—	C	C	13	3
7.	—	—	—	—	C	C	D	C	D	C	C	C	C	C	D	C	—	C	C	—	11	3
8.	C	—	D	D	—	D	D	D	D	C	D	C	D	D	D	C	D	C	D	D	5	13
9.	C	C	—	—	C	D	D	C	C	D	C	D	D	D	D	D	D	—	D	D	6	11
10.	—	—	—	D	D	D	D	D	C	C	C	D	C	D	C	C	—	—	C	D	8	7
11.	—	—	—	—	D	C	—	C	C	D	C	D	C	C	—	—	C	C	C	C	10	3
12.	D	C	D	C	—	—	D	C	—	C	C	D	D	D	D	D	D	C	D	C	6	11
13.	D	D	C	D	D	D	D	D	C	C	C	D	C	C	C	C	C	D	C	D	10	10
14.	D	—	—	C	—	—	—	—	C	D	C	C	D	D	—	—	—	—	—	—	4	4
15.	D	—	C	—	—	—	—	—	D	—	C	C	D	C	C	C	D	C	C	C	9	4
16.	—	—	—	C	—	C	D	D	D	C	D	—	—	C	D	C	—	C	D	D	6	7
17.	D	D	C	C	—	—	—	C	C	C	C	D	C	—	C	C	C	D	D	C	11	5
18.	—	D	—	—	—	—	D	C	C	D	—	—	D	C	C	—	C	—	D	D	5	6
19.	C	—	C	C	C	C	D	C	C	C	C	D	—	—	—	C	—	—	C	D	11	3
20.	C	—	C	—	D	—	—	—	D	C	C	D	C	D	C	D	—	—	D	—	7	5
21.	C	C	C	D	D	D	C	D	C	C	D	C	D	D	D	D	D	C	D	C	9	11
22.	—	—	—	—	—	—	—	—	—	—	—	—	—	—	D	—	—	—	—	—	0	1
23.	D	D	D	C	C	C	C	—	—	D	—	—	C	C	—	—	D	C	—	C	8	5
24.	D	D	D	—	C	C	C	C	—	—	D	—	—	—	—	—	C	C	D	D	6	6
25.	—	—	D	—	—	—	C	—	—	D	C	C	C	D	C	C	C	C	C	C	10	3
26.	C	D	—	C	D	C	D	C	D	C	D	D	C	—	D	D	D	D	C	D	7	11
27.	C	D	C	—	D	D	D	D	D	—	C	D	D	D	—	—	D	—	C	D	4	11
28.	D	D	D	C	D	C	D	C	D	C	D	C	C	C	D	C	C	C	D	D	10	10
29.	C	C	C	C	C	D	D	C	—	—	—	—	—	—	—	—	—	—	D	—	6	3
30.	D	D	D	D	D	D	C	C	—	—	—	—	D	C	D	C	C	D	C	C	7	9

Totals for the Ten Days

Chick to chick (C)	235 times	39.1%
Chick to duck (D)	193 times	32.1%
Chick to neither (—)	172 times	28.7%

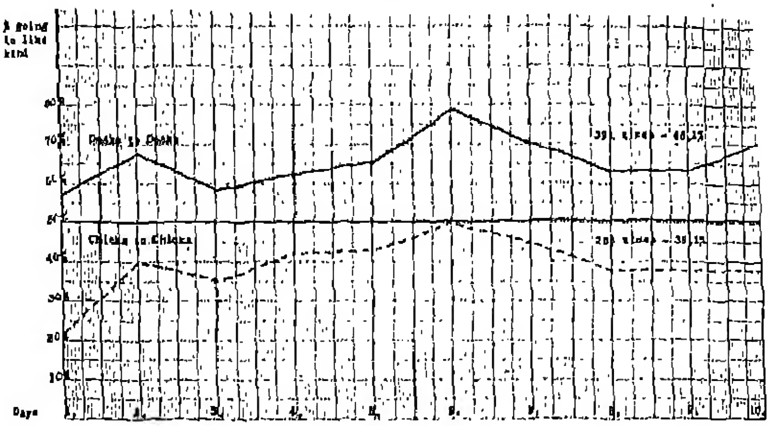

FIGURE 3

SHOWING RELATIVE STRENGTH OF TENDENCY OF CHICKS AND DUCKS TO GO TO LIKE KIND—IN CHOICE CAGE

Each curve represents 60 daily trials for each kind of animal—two each for 30 chicks and 30 ducks, respectively—a total of 600 trials for each curve. Experiments were conducted daily for the first 10 days after hatching.

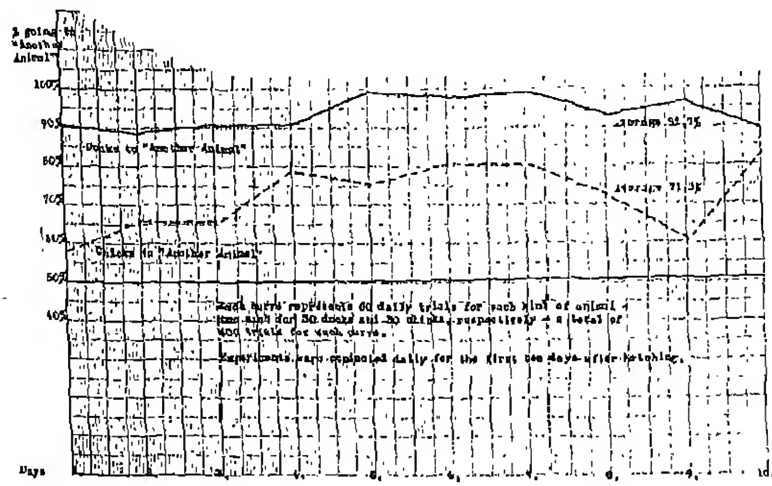

FIGURE 4

GRAPH SHOWING RELATIVE STRENGTH OF TENDENCY OF CHICKS AND DUCKS TO GO TO "ANOTHER ANIMAL," EITHER CHICK OR DUCK—IN CHOICE CAGE

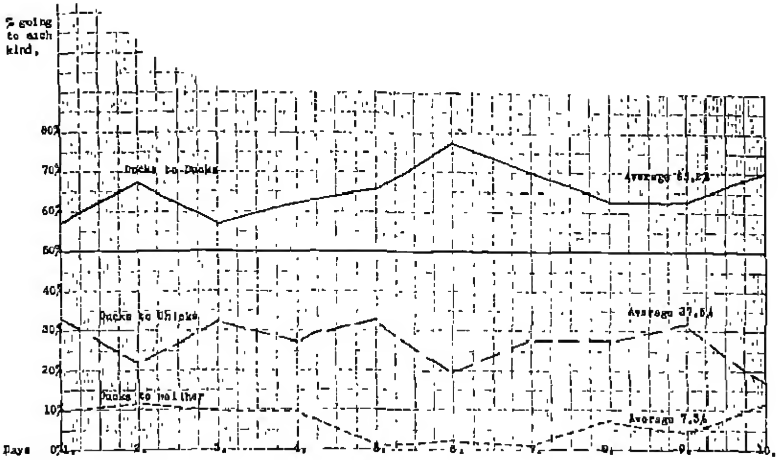

FIGURE 5

GRAPH SHOWING RELATIVE STRENGTH OF TENDENCY OF DUCKS TO GO TO DUCKS, TO CHICKS, AND TO NEITHER ANIMAL—IN CHOICE CAGE

Each curve represents 60 daily trials for each kind of animal—two each for 30 ducks and 30 chicks, respectively—a total of 600 trials for each curve—for first 10 days after hatching.

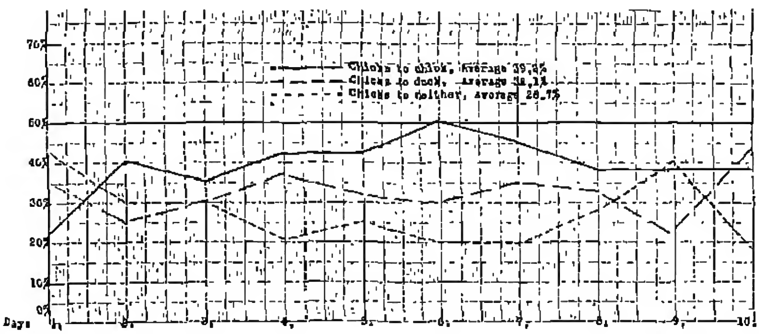

FIGURE 6

GRAPH SHOWING RELATIVE STRENGTH OF TENDENCY OF CHICKS TO GO TO CHICKS, TO DUCKS, AND TO NEITHER ANIMAL—IN CHOICE CAGE

Each curve represents 60 daily trials for each kind of animal—two each for 30 chicks and 30 ducks, respectively—a total of 600 trials for each curve for the first 10 days after hatching.

behavior appears warrants the assumption of a "predisposition" in favor of certain types of behavior, in harmony with other physiological and psychological aspects of organization.

4. That we must regard psychological inheritance (i.e., psychological tendencies and abilities) as an aspect of the total integrated pattern of innate constitution.

With the demonstration of psychological inheritance at a given point in the phylogenetic scale it seems highly probable that other general tendencies and abilities may constitute at least an innate potential aspect of not only animal but also of human behavior.

6. *Comment*

The more qualitative aspects of the observed chick and duck behavior merit comment. The behavior of the duck, for example, is far more impulsive than that of the chick. The duck ordinarily makes his approach in about one-half the time consumed by the chick. After the 30-second period allowed for orientation the duck invariably runs headlong from the cage "*C*" toward the animal of its "choice." Frequently upon reaching the other duck it will immediately lie down as close to the other duck as possible, uttering a series of peculiar, rapid, subdued sounds which are heard at no other time. This behavior is frequently reciprocated by the other duck, who in turn lies as closely as possible, uttering the same sounds. In only one instance have I observed the released duck exhibit this behavior toward a chick. This qualitative difference in behavior may be regarded as the affective aspect of the total innate organization of "duck." It is part of what makes a duck a duck. Impulsiveness may, perhaps, account in part for the duck's stronger tendency to "approach another animal." But impulsiveness, as such, cannot account for the greater discrimination of kind or the stronger tendency and ability to go to own kind.

There appears to be some basis for early differentiation of sex behavior. A most promising suggestion for investigating the early appearance of sex differentiation lies in the difference in behavior of the little *ducks* and little *drakes*. Whereas the former rush precipitately to another animal, the former stand erect and elert upon hasty egress from the tunnel, approaching neither animal or only after a period of "observation." Here is apparently the germ of differential sex organization appearing at a very early date. This discovery was made when, as grown ducks, certain wing-banded

animals were found to be leading the flock, and too late to identify all of them. Nos. 5 and 16 in Table 2 are drakes, and together account for 25 per cent of the failure to go to either animal in the group of 30 ducks used in the experiment.

B. Variations of Experiment

Another group of chicks and ducks (13 each) was hatched, at the same time, in the same incubator, and in equal numbers of each, as in Experiment I. These animals were used in a variety of modifications of Experiment I, using the same apparatus (see Figure 2). The experimental technique employed was similar throughout, the trials covering a period of 15 days, beginning the third day after hatching. The nature of these variations together with the results for each are summarized briefly, in terms of approach to like or unlike animal or neither animal.

1. *A Control Experiment*

The 13 animals of each group were put through a series of trials identical with that of Experiment I (*a*) as a check on Experiment I, (*b*) as a basis for comparison with performance in the modified experiments.

Results

Duck to duck	72.0%	Chick to chick	51.0%
Duck to chick	27.5%	Chick to duck	39.4%
Duck to neither	.5%	Chick to neither	9.6%
(204 trials)		(208 trials)	

Results for these groups are substantially the same as for the groups used in Experiment I.

2. *Pointing Cage "C" Toward Like Animal*

What would be the effect of pointing Cage "C" toward *Like* animal, i.e., of deliberately favoring the approach of each animal toward *Like* kind? In each trial, starting cage "C" was pointed toward the cue animal corresponding to the one about to be placed in starting cage.

Results

Duck to duck	71.0%	Chick to chick	50.5%
Duck to chick	29.0%	Chick to duck	40.8%
Duck to neither	.0%	Chick to neither	8.7%
(103 trials)		(103 trials)	

Apparently "weighting" the experiment in favor of going to like

kind had little effect upon choice, though it did influence "approach."

3. *Pointing Cage "C" Toward Unlike Animal*

What would be the effect of deliberately *favoring* the approach of each animal toward *Unlike* kind?

Results

Duck to duck	59.0%	Chick to chick	41.7%
Duck to chick	40.0%	Chick to duck	43.7%
Duck to neither	1.0%	Chick to neither	14.6%
(102 trials)		(104 trials)	

Pointing the cage toward *unlike* animal definitely affected the choice of animal approached, with both ducks and chicks, decreasing the number of approaches to like kind. In spite of this unfavorable "weighting," however, the ducks still tended definitely to approach like kind, usually sweeping rapidly in a wide curve toward their own kind.

4. *With Duck in One Cage, Other Cage Empty*

In this instance there was no *choice* of animal to be approached. Cage *"A"* or *"B"* contained one duck while the other cage was empty. What is the strength of the (gregarious) tendency to approach "another animal"?

Results

Duck to "other animal," a duck 90.0% (100 trials)
Chick to "other animal," a duck 59.0% (100 trials)

5. *With Chick in One Cage, Other Cage Empty*

Results

Duck to "other animal," a chick 87.0% (100 trials)
Chick to "other animal," a chick 63.0% (100 trials)

The tendency to approach "another animal" is about the same as when a choice is offered, with a slightly stronger tendency to approach when the animal is of *like* kind.

6. *Animals in Starting Cage "C" Can Hear But Not See Cue Animals*

In this experiment a modification of the apparatus was used (see Figure 7.) A heavy cloth was hung over the near sides of Cages *"A"* and *"B,"* obscuring cue animals from view for the animal in Cage *"C."* Other than visual cues must now be employed in approaching like or unlike kind.

Results

Duck to duck	77.0%	Chick to chick	39.4%
Duck to chick	22.0%	Chick to duck	15.6%
Duck to neither	1.0%	Chick to neither	45.0%
(108 trials)		(109 trials)	

The much greater deliberation and uncertainty with which the ducks now approach other animals, in contrast with the alacrity and impulsiveness which they exhibited in the previous experiments, suggests that, while there appears to be good auditory discrimination, vision plays a greater if not more important rôle in this type of experimentation. Chicks, when they approach another animal, still tend definitely to approach like kind, but their "indifference" when they lack visual cues is very pronounced, and in marked contrast to the behavior of the ducks. The chicks approached "neither" in 45 per cent of the trials.

A qualitative difference is also suggested. Hearing the cries of other animals, while unable to see them, appears to have a more profound influence on the ducks than on the chicks. The ducks' behavior upon release from "*C*" appears much more "concerned" and uncertain. This qualitative difference is borne out even more distinctly in later experiments. Another qualitative difference in the behavior of ducks and chicks is observable in the fact that the ducks travelled a greater distance toward like kind, upon release from "*C*," roughly comparable to the relative strength of ducks and chicks to approach like kind.

Duck

Travelled to duck 476 in.—78.4% of total distance travelled.
Travelled to chick 131 in.—21.6% of total distance travelled.

Chick

Travelled to chick 280 in.—57.0% of total distance travelled.
Travelled to duck 212 in.—43.0% of total distance travelled.

C. Experiment II: Tendency of Newly-Hatched Chicks and Ducks to Approach a Group of Like or Unlike Kind

1. *Problem*

How will chicks and ducks (hatched at the same time, in the same incubator, in equal numbers of each, and reared together in the same environment from birth, thus with "equal opportunity to learn") respond to the presence of *groups* of animals of like and unlike kind upon release from confinement?

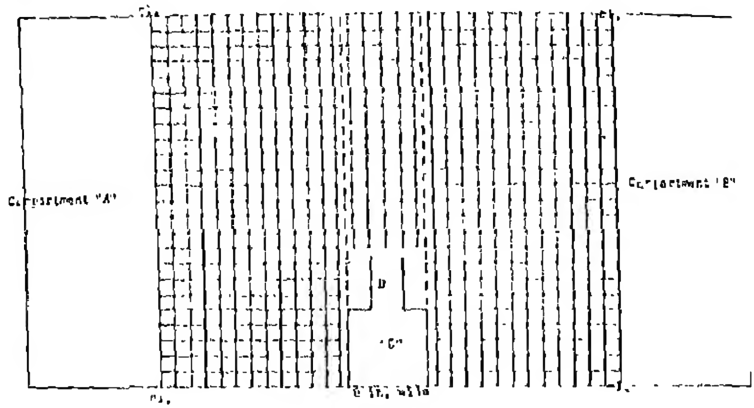

FIGURE 7

Choice Cage

Experiment II: Choice in going to like or unlike kind—in a situation in which the animals cannot see one another.

A duck and chick are placed in "*A*" and "*B*" respectively (alternated with each series). White flannel cloth "*cl*" covers the side wall of compartments "*A*" and "*B*."

Animals are released singly from "*C*" through tunnel in response to cries and their behavior noted with respect to movements toward either compartment, as they pass the broken lines in either direction, toward like or unlike kind.

The floor is marked off in inch squares to facilitate measurement in any direction.

2. *Materials and Procedure*

A third group of animals, 19 Buff Orpington chicks and 19 White Pekin ducks, were hatched as in the other experiments. The apparatus used is illustrated and described in Figure 8. After a 30-second period of orientation single animals were released from starting Cage "*C*" and behavior was noted with respect to approach to *groups* of like and unlike kind.

These animals were reared in two separate groups: one group consisting of 12 chicks and 12 ducks, the other of 7 chicks and 7 ducks. Animals of one group were used as cue animals for the other group, and vice versa. The only time animals of the two groups saw one another was during the actual experimentation. None of these animals were used in any other experiment. The equal numbers of unlike animals were kept in close confinement from hatching, and experiments were begun when the animals were 18 days old.

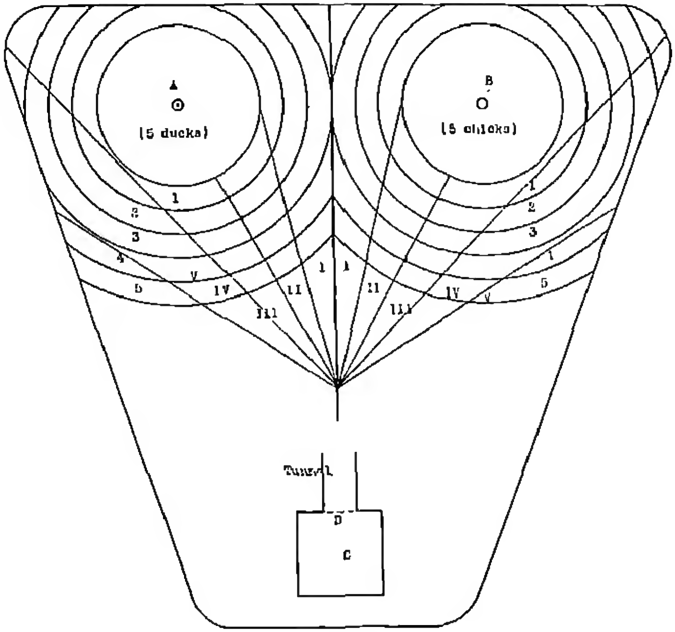

FIGURE 8

CHOICE OF CHICKS AND DUCKS IN GOING TO A GROUP OF CHICKS OR DUCKS

Animals are released through sliding door *D* raised by a string over a pulley on a tripod overhead.

Distance from center of *A* or *B* to end of tunnel is 4 ft.

Animals are released from Cage *C*.

A and *B* contain 5 ducks or 5 chicks, alternating with each series.

Spaces numbered 1-5 and I-V are designed for a quantitative analysis if desired.

Base of apparatus is made of inverted linoleum. Lines on apparatus are barely discernible.

A and *B* enclosed with 1 ft. ½ in. wire mesh. The whole is enclosed with 2 ft. mesh.

Results

Duck approached duck: 240 times or in 83.9% of trials.
Duck approached chick: 26 times or in 9.1% of trials.
Duck approached neither: 20 times or in 7.0% of trials.
(Total of 286 trials)

Chick approached chick: 145 times or in 50.7% of trials.
Chick approached duck: 111 times or in 38.9% of trials.
Chick approached neither: 30 times or in 10.4% of trials.
(Total of 286 trials)

The tendency of chicks to approach *groups* of like kind is substantially the same as with the approach to single animals in previous experiments. The slight accentuation in choice may be due to increased age. During the rather extended period (18 days) of this equivalent preliminary confinement, the animals did tend to segregate more or less into homogeneous groups. Instead of militating against the experiment, however, this fact may be taken to indicate that, within the equivalent environmental setting, differentiating factors were already at work. This fact can but serve to strengthen, therefore, the conclusion that the various aspects of structural and functional differences which contribute to early homogeneous segregation are also closely integrated with constitutional differences in tendencies and abilities.

3. *Conclusion*

This gregarious tendency of chicks and ducks to associate with "other animals," and especially with like kind, and, as Schjelderup-Ebbe says, the ability to "learn more easily" to go to like rather than to unlike kind, established and differentiated against a background of equivalent environmental experience from birth, indicates that somehow this tendency and discriminatory ability are rooted in, and are aspects of innate constitution, along with their structuro-functional differentiation. They lend support to McDougall's contention that the cognitive, effective, and conative tendencies and abilities of animals are rooted in innate constitution. Further, it seems reasonable to suppose that this same psycho-biological integration may be ascribed to all psychological behavior, of both men and animals.

IV. EXPERIMENTAL STUDIES: SOCIAL BEHAVIOR IN RESPONSE TO CRIES

A. Experiment II: Affective Responses to Shocking of Like or Unlike Animals, Under Equivalent Environmental Conditions

The aim of this experiment was twofold. (*a*) On the one hand it is substantially the same as in the foregoing experiment, namely, to determine whether there is any innate basis for differences in cognitive abilities and behavior tendencies in chicks and ducks. In this instance we are interested, however, in possible early differentiation of response to cries of like and unlike kind. (*b*) On the other hand, it embodies an attempt to equalize the factors in the actual experimental set-up as well as in the general environmental setting prior to experimentation.

It might be thought, in the earlier "Choice" situations, that perhaps the chicks and ducks were influenced in their "choice" by other factors than those differentiating like and unlike kind; that, for example, they might be responding to slight differences in size, color, intensity of chirp, etc., rather than to "like or unlike kind." It might be theoretically possible to equate some of these factors individually, to secure animals of equal size, exact similarity of color, equally loud chirp, etc. But, even though these difficulties were surmounted, there might be other factors, either unnoticed or unthought of, which might still play an important part in influencing behavior. It would be desirable, therefore, to provide an experimental situation in which all extraneous factors might be eliminated or at least equated, simultaneously. For this purpose a new and different technique is suggested to provide not only equivalent experience prior to the experiment, but to maintain this equivalence in the experimental situation itself.

1. *The Problem*

How will chicks and ducks (hatched together, and reared from birth in the same environment, in small groups of equal numbers of each) respond to cries of pain or alarm from like and unlike animals? A special feature is an environmental set-up designed to equalize all the possible factors influencing behavior—except that constellation of factors which differentiate the total integrated "nature" of chicks and ducks.

2. *Materials*

An "octagonal cage" was constructed on a wood base, 30 inches in width. This was divided into eight compartments so that the animals might be placed in them in the order, chick, chick, duck, duck (see Figure 9). Thus each chick and each duck was flanked by both "like" and "unlike" kind.

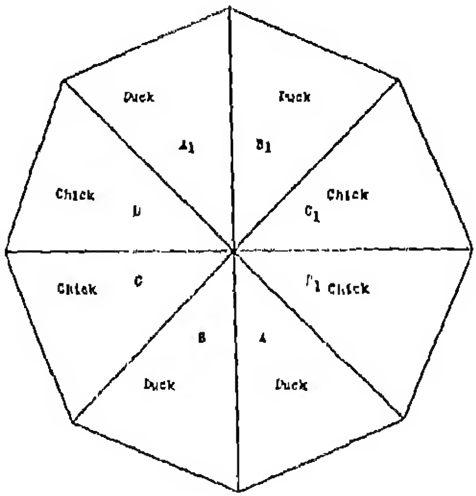

FIGURE 9

OCTAGON CAGE USED IN EXPERIMENTS IV AND V

Chicks and ducks are placed in compartments, as shown. Each animal is thus flanked by an animal of like and unlike kind. They may all be shocked simultaneously or in alternate pairs, of *A*'s, *B*'s, *C*'s, and *D*'s.

The floor of each compartment is furnished with removable electric shocking grid, wired for simultaneous or alternate shocking. Partitions are of ½ inch wire mesh; the outer edge of pasteboard.

A motion-picture camera, suspended overhead on a tripod, photographs all movements and positions simultaneously.

This method and procedure were suggested by Dr. T. H. Howells, of the University of Colorado, as a technique for determining the effects of "equivalent conditioning."

The compartments were separated from one another by one-half inch wire mesh, so that the animals were easily visible to one another. The front of the cage was enclosed with heavy paste-board, while the center-post was surrounded with opaque glass, so the animal's vision was restricted to either side of its compartment. The floor

of each compartment was furnished with removable electric shocking grids. These grids were wired in such a manner and equipped with separate switches so that a shock might be given to any pair of like animals (the *A*'s, *B*'s, *C*'s, or *D*'s), or to all the animals simultaneously. The shocking apparatus itself consisted of a battery of eight (Model T Ford) coils and a transformer, using the ordinary house current.

Eight chicks and eight ducks were used. (These animals were not used in any other experiment.) They were used in two groups of four chicks and four ducks each, to meet the capacity of the octagonal cage. Experiments were carried on successively for the groups on the same days.

3. *Procedure*

Animals were placed in the compartments of the octagonal cage daily, from the 10th to the 20th day after hatching, in the order chick, chick, duck, duck, as shown in Figure 9.

Upon being placed in the apparatus, the animals, especially the ducks, showed a tendency to approach an animal in one of the adjacent compartments. The animal approached was frequently one of like kind. (This is doubtless an expression of the differentiation already going on within the equivalent environment, preceding the experiment, and must therefore be regarded as a part of the total differentiating process.) In a very short time (usually less than one or two minutes), the ducks would lie down, while the chicks would remain standing quietly. This was taken as a signal for giving the shock. In each instance the shock was given 15 seconds after the last of the four ducks had lain down.

Chicks and ducks were given shocks of approximately one and three thousand milliamperes respectively (as calculated, the difference in shock being relative to their sensitivity to shock). Glycerine, with an admixture of soda and water was used to moisten the feet of the chicks so that a shock of lower intensity might be used than would otherwise be required.

Each daily series of shocks involved three steps, as follows: (*a*) Alternate pairs of animals were shocked, proceeding clockwise, in the order, *A*'s, *B*'s, *C*'s, *D*'s. There was a two-minute interval between the shocking of each pair. One shock of about one second's duration was given. (*b*) All animals were shocked simultaneously—a series of five shocks of one second each, at 15-second intervals.

(*c*) Alternate pairs were again shocked, as before, in order to note the difference, if any, produced by receiving the shock at the same time that the cries of the others were heard.

4. *Recording*

A motion-picture camera was suspended on a tripod directly overhead, recording both *positions* and *movements* of all animals simultaneously. This recording was begun just before shocking and was continued for some 25 (16 mm.) frames. Some 200 feet of motion-picture film was used for this experiment, which could be studied at leisure to supplement the customary data of direct observation.

Observation was of two kinds:

1. The *movements* of the animals. Interest centers, not about the movements of the shocked animals but about the movements of the animals (like and unlike) adjacent to the shocked animal on either side. What difference, if any, may be noted with respect to the responses of like and unlike kind?

2. The *position* taken by each animal in each instance, before and after, with reference to like and unlike animals. These positions were noted daily, as follows: (*a*) Just before the first shock was administered (ducks lying), (*b*) immediately after shocking, (*c*) immediately after simultaneous shocking, (*d*) just before re-shocking (ducks lying), (*e*) immediately after re-shocking.

5. *Interpretation*

In summating the responses to like and unlike animals, it is important to appreciate the fact that any pronounced or excessive stimulating effect that the behavior of a given kind of animal might have, would result in an increase (if this were a determining factor in behavior) in response not only of an adjacent animal of his own kind but would also increase the response of an unlike animal on the other side.

It might be argued, for example, that chicks make more noise than ducks, have a higher pitched voice, are less susceptible to shock, etc. But it is again a pointed advantage of the present method that it automatically eliminates such extraneous factors.

For the increment in response resulting from such an excessive stimulus (e.g., the louder noise made by the chick) would increase the response not only of a like animal (another chick), but also of an unlike animal (e.g., a duck). In the first instance, the increment

would be added, of course, to the sum of responses to *like* animals. In the second instance it would increase the sum of responses to *unlike* animals. Thus these increments would tend to cancel each other out insofar as their effect on like or unlike responses is concerned.

To illustrate, let us suppose that chicks and ducks make sounds of equal volume. Then, obviously, loudness of sound would not constitute a determining factor in their response.

But, suppose chicks do make a louder sound than do the ducks. Then it might be supposed that chicks and ducks, instead of responding to cries of like or unlike kind, are responding rather to this loudness of sound, or at least that this is an important factor. Yet, this is not the case under the "double" arrangements, for, let us say that chicks and ducks normally respond with 50 units of response to the strength of sound normally emitted by a duck. And let us say that the greater loudness of noise produced by chicks increases the response by 50 additional units. This would give us the following results (Table *A*) proceeding clockwise around the cage.

TABLE *A*

	Units of response to like	Units of response to unlike
Let Chick "C" be shocked		
Then Chick "*D*" would make 100 units of response to sound	100	
Then Duck "*B*" would make 100 units of response to sound		100
Let Chick "D" be shocked		
Then Chick "*C*" would make 100 units of response to sound	100	
Then Duck "*A*" would make 100 units of response to sound		100
Let Duck "A" be shocked		
Then Chick "*D*" would make 50 units of response to sound		50
Then Duck "*B*" would make 50 units of response to sound	50	
Let Duck "B" be shocked		
Then Chick "*C*" would make 50 units of response to sound		50
Then Duck "*A*" would make 50 units of response to sound	50	
Total:	300 L.	300 U.

The result would be 300 units of response to like kind and 300 units of response to unlike kind. Obviously the animals are responding to something other than "like and unlike kind," and this tends to cancel itself out automatically. The same would hold true for differences in sensitivity to shock, for differences in pitch of voice, size, color, etc.

Not only does this hold true, but any other extraneous factors, which might not be at all anticipated, or which might even be entirely unknown, would thus tend to cancel themselves out.

All factors, except those which differentiate *responses to like and unlike kind* would tend to cancel themselves out automatically. And only those factors, which differentiate specific behavior towards *like and unlike* would remain—just those factors whose existence we are attempting to investigate. This method, therefore, affords a unique method for equating environmental stimuli and at the same time permits a differentiation of factors distinguishing responses between like and unlike animals.

How then may this method be applied to differentiate responses

TABLE B

	Units of response to like	Units of response to unlike
Let Chick "C" be shocked		
Then Chick "*D*" would make 50 units of response to cries	50	
Then Duck "*B*" would make 0 units of response to cries		0
Let Chick "D" be shocked		
Then Chick "*C*" would make 50 units of response to cries	50	
Then Duck "*A*" would make 0 units of response to cries		0
Let Duck "A" be shocked		
Then Chick "*D*" would make 0 units of response to cries		0
Then Duck "*B*" would make 50 units of response to cries	50	
Let Duck "B" be shocked		
Then Chick "*C*" would make 0 units of response to cries		0
Then Duck "*A*" would make 50 units of response to cries	50	
Total:	200 L.	0 U.

between like and unlike kind? Let us suppose that animals *do* respond only to the cries or other behavior of *like* kind. Then we would obtain the following results, in proportion as they do consistently respond (Table *B*).

The result would then be a more or less consistent response of animals to *like*—hence a discriminating response, in proportion to the consistency of the behavior.

6. *Results*

a. Amount of movement, in response to shocking of adjacent like and unlike animals. In order to obtain some estimate of the degree of response, in terms of amount of bodily movement, the motion pictures were thrown on a screen so that they might be studied, one frame at a time. The projector was placed five feet from the screen, making an image 10 inches across. Amount of movement was then measured in terms of change of position from the first to

TABLE *C*

Upon shocking ducks	
Adjacent ducks moved *more* than adjacent chicks	51 times or 63.8% of trials.
Adjacent ducks moved *less* than adjacent chicks	24 times or 30.0% of trials.
Adjacent ducks moved *same* amount as adjacent chicks	5 times or 6.2% of trials.
Upon shocking chicks	
Adjacent chicks moved *more* than adjacent ducks	48 times or 60.0% of trials.
Adjacent chicks moved *less* than adjacent ducks	25 times or 31.2% of trials.
Adjacent chicks moved *same* amount as adjacent ducks	7 times or 8.8% of trials.
Upon re-shocking ducks	
Adjacent ducks moved *more* than adjacent chicks	48 times or 60.0% of trials.
Adjacent ducks moved *less* than adjacent chicks	28 times or 35.0% of trials.
Adjacent ducks moved *same* amount as adjacent chicks	4 times or 5.0% of trials.
Upon re-shocking chicks	
Adjacent chicks moved *more* than adjacent ducks	49 times or 61.3% of trials.
Adjacent chicks moved *less* than adjacent ducks	19 times or 23.7% of trials.
Adjacent chicks moved *same* amount as adjacent ducks	12 times or 15.0% of trials.

the twentieth frame, using the millimeter as a unit and the point of the beak as the point of reference.

The 10 daily series of shocking experiments with eight chicks and eight ducks gave a total of 80 measurements for each. The results, for these 80 trials, are given in Table *C*.

Both chicks and ducks moved considerably *more* when *like* kind were shocked than when *unlike* kind were shocked.

b. Direction of movement, position taken by animals relative to like and unlike kind. In order to obtain some estimate of response in terms of *direction* of movement, the motion pictures were again thrown on the screen, one frame at a time. Direction of movement was then determined in terms of *change of position*, relative to like and unlike kind, by noting the direction of movement as shown from the first to the twentieth frame. If an animal maintained his position relative to a given adjacent animal, or moved to a position on one side of the compartment and faced the animal on that side

TABLE *D*

Positions taken by ducks adjacent to shocked pairs of animals

Before shocking adjacent ducks: Duck to duck 56 times or 70.0% of trials.
After shocking adjacent ducks: Duck to duck 44 times or 55.5% of trials.
Before re-shocking adjacent ducks:
Duck to duck 63 times or 78.8% of trials.
After re-shocking adjacent ducks: Duck to duck 57 times or 71.3% of trials.
Before shocking adjacent chicks: Duck to duck 61 times or 76.3% of trials.
After shocking adjacent chicks: Duck to duck 63 times or 78.8% of trials.
Before re-shocking adjacent chicks:
Duck to duck 69 times or 86.3% of trials.
After re-shocking adjacent chicks:
Duck to duck 65 times or 81.3% of trials.

Positions taken by chicks adjacent to shocked pairs of animals

Before shocking adjacent chicks: Chick to chick 52 times or 65.0% of trials.
After shocking adjacent chicks: Chick to chick 55 times or 68.8% of trials.
Before re-shocking adjacent chicks:
Chick to chick 52 times or 65.0% of trials.
After re-shocking adjacent chicks:
Chick to chick 54 times or 67.5% of trials.
Before shocking adjacent ducks: Chick to chick 50 times or 62.5% of trials.
After shocking adjacent ducks: Chick to chick 56 times or 70.0% of trials.
Before re-shocking adjacent ducks:
Chick to chick 55 times or 68.8% of trials.
After re-shocking adjacent ducks:
Chick to chick 56 times or 70.0% of trials.

Positions taken when all animals were shocked simultaneously

Duck to duck 64 times or 80.0% of trials.
Chick to chick 63 times or 78.8% of trials.

rather than the animal on the other side, it was regarded as inclining toward that animal. When it was difficult to determine the position relative to adjacent animals it was regarded as "indifferent" as to position. In the large majority of cases the tendency to "approach another animal" was very pronounced.

The 10 daily series of shocking experiments with eight chicks and eight ducks gave a total of 80 recordings for each kind of animal. The results may be summarized as in Table *D*.

7. *Conclusion*

Unfortunately, this experiment, designed to investigate differentiated response to cries of like and unlike kind, affords little or no evidence for such response.

For the ducks, it indicates the same tendency to approach like kind as was found in Experiment I. It is not clear whether the *cries* of the adjacent shocked animals was an important factor in this gregarious tendency. It will be noted, however, that when the adjacent ducks were shocked the other ducks exhibited less tendency to approach like kind, especially after the first shocking of the daily series, their behavior being characterized in some instances by violent attempts to escape from the cage. With the re-shocking of the adjacent ducks, however, this escape behavior was less pronounced, and the gregarious tendency somewhat more in evidence. Shocking of the adjacent chicks on the other hand, was characterized by fewer attempts to escape from the cage and by a somewhat greater tendency to move toward the other ducks.

For the chicks, the results indicate a stronger tendency on the part of the chicks to approach like kind than was found in Experiment I. Again, it is not clear whether the *cries* of the adjacent shocked animals was an important factor in initiating the response. The chicks, however, showed little or no difference in their behavior, either before or after shock, or when adjacent animals of like or unlike kind were shocked.

In addition to these slight differences, the equivalent experimental conditions of the peculiar arrangement probably serve to emphasize that the animals were responding to "like and unlike kind" rather than to accidental minor differences in color, size, intensity of chirp, etc.

B. Experiment III: Affective Responses to Cries of Like and Unlike Animals in the Dark

As both chicks and ducks showed considerable "distress" upon being isolated from one another, especially in the dark, when they could hear but neither see nor contact one another, it was decided to take advantage of this fact in an experimental situation. This had the added advantage of a more or less "natural" situation rather than the more artificial one of electric stimulation.

1. *The Problem*

How will chicks and ducks respond to cries of like and unlike animals when isolated from one another in darkness?

2. *Materials*

The same octagonal cage (Figure 9) was used as in the previous experiment except that the electric grids were removed from each of the eight compartments. Thus each chick and each duck was flanked by both "like" and "unlike" kind.

Twenty-eight chicks and 28 ducks were used in this experiment in order that we might have seven sets of four chicks and four ducks each, to meet the requirements of the octagonal cage. The animals were hatched together, of course, in equal numbers of each, and reared together in close confinement.

3. *Procedure*

The animals (which were banded and numbered, as usual) were placed in their respective compartments in the octagonal cage, in the dark, in a dark room. There they were left for a period of six minutes, during which time the chirping was loud and incessant. As this experiment was designed particularly to study the behavior of animals with respect to cries from adjacent animals it was decided to note the positional responses both in the dark and immediately after switching on the light.

(*a*). Position of the animals was noted at the end of the six-minute period of darkness, before the light was switched on. A very weak beam of light from a pocket flash-light was cast momentarily upon the ceiling of the room at the end of the first five minutes of the six-minute period. This light was just strong enough to enable the two observers to note the relative positions of the animals

in the cages (the positions of the cages being noted beforehand), though scarcely strong enough to recognize the animals.

(*b*). At the end of the six-minute period the light was switched on. Fifteen seconds later the positions of all the animals, with respect to their approach to like or unlike kind, was recorded simultaneously by two observers who had been seated quietly throughout the entire period a short distance away, from a position where they could look down into the cage. Results are given in Table *E*.

TABLE *E* (*in darkness*)

(28 ducks—1 trial daily for ten days—total of 280 trials)	
Duck approached duck	153 times or 54.6% of 280 trials
Duck approached chick	94 times or 33.6% of 280 trials
Duck approached neither	33 times or 11.8% of 280 trials
(28 chicks—1 trial daily for ten days—total of 280 trials)	
Chick approached chick	129 times or 46.1% of 280 trials
Chick approached duck	121 times or 43.2% of 280 trials
Chick approached neither	30 times or 11.7% of 280 trials

As this period of darkness was characterized by much loud chirping and attempts to escape from the cage, these figures cannot be regarded as a measurement of the tendency to approach another animal nor the ability to discriminate like and unlike kind. Yet the figures indicate some tendency to approach like rather than unlike. *When the lights were switched on one minute later, attempts to* escape were largely succeeded by approach to other animals.

Immediately upon switching on the light the cries of the animals stopped abruptly and almost completely. Those who were not already close together almost invariably ran quickly to one or another of the adjacent animals. Approach to another animal was usually as close as the cages permitted, especially in the case of the ducks, and the positions were recorded with respect to that animal. In the great majority of cases the animal took up a definite position with respect to one or another of the adjacent animals. In a small number of cases the position of a given animal was determined as toward chick or duck according as its head was to one or other side of the cage line dividing its cage horizontally. In a few instances, when there was doubt, or the animal's body ranged horizontally along the dividing line, the animal was regarded as "indifferent" (Table *F*).

TABLE F (*after turning on light*)

(*28 ducks—1 daily trial for ten days—total of 280 trials*)	
Duck approached duck	196 times or in 70.0% of the trials.
Duck approached chick	70 times or in 21.4% of the trials.
Duck approached neither	14 times or in 8.6% of the trials.
(*28 chicks—1 daily trial for ten days—total of 280 trials*)	
Chick approached chick	145 times or in 51.8% of the trials.
Chick approached duck	89 times or in 31.8% of the trials.
Chick approached neither	46 times or in 16.4% of the trials.
Incidentally it will be noted from the above figures that:	
Ducks approached "another animal"	266 times or in 91.4% of the trials
Chicks approached "another animal"	234 times or in 83.6% of the trials

4. *Conclusion*

The ducks display an ability to discriminate the cries of like and unlike kind in the dark, the results being closely comparable to those obtained in Experiment I, *F*, in which visual perception of cue animals was eliminated. The behavior of the chicks, however, was undifferentiated in this experiment, with respect to like and unlike kind. In addition, the behavior of the chicks was quite different from that in Experiment I, *F*, perhaps because *both* experimental and cue animals were in the dark. This suggests a differentiation in chicks and ducks both with respect to visual ability and emotional response to darkness.

TABLE 4

SHOWING THE POSITIONAL RESPONSES OF CHICKS AND DUCKS
In the dark—at the end of the first five minutes of the six-minute period of darkness.

Days	Duck's behavior: Duck to duck Number and % of cases		Duck's behavior: Duck to chick Number and % of cases		Chick's behavior: Chick to chick Number and % of cases		Chick's behavior: Chick to duck Number and % of cases	
1.	15	53.6%	10	35.7%	15	53.6%	12	42.9%
2.	16	57.1%	11	39.3%	16	57.1%	10	35.7%
3.	12	42.9%	13	46.4%	13	46.4%	14	50.0%
4.	21	75.0%	6	21.4%	12	42.9%	13	46.4%
5.	15	53.6%	12	42.0%	14	50.0%	11	39.3%
6.	10	35.7%	17	60.7%	11	39.3%	13	46.4%
7.	21	75.0%	7	25.0%	12	42.9%	14	50.0%
8.	16	57.1%	11	39.3%	13	46.4%	10	35.7%
9.	16	57.1%	9	32.1%	14	50.0%	10	35.7%
10.	17	60.7%	8	28.6%	9	32.1%	14	50.0%
Totals	153		94		129		121	
Averages		54.6%		33.6%		46.1%		43.2%

TABLE 5

SHOWING THE POSITIONAL RESPONSES OF CHICKS AND DUCKS

Fifteen seconds after the light had been turned on following the six-minute period of darkness.

	Duck's behavior				Chick's behavior			
	Duck to duck Number and % of cases		Duck to chick Number and % of cases		Chick to chick Number and % of cases		Chick to duck Number and % of cases	
Days								
1.	18	64.3%	10	35.7%	15	53.6%	11	39.3%
2.	18	64.3%	10	35.7%	13	46.4%	11	39.3%
3.	19	67.9%	6	21.4%	18	64.3%	6	21.4%
4.	16	57.1%	11	39.3%	15	53.6%	8	28.6%
5.	19	67.9%	4	14.3%	11	39.3%	12	42.9%
6.	18	64.3%	10	35.0%	15	53.6%	10	35.7%
7.	19	67.9%	8	28.6%	17	60.7%	7	25.0%
8.	21	75.0%	4	14.3%	16	57.1%	8	28.6%
9.	23	82.1%	4	14.3%	17	60.7%	7	25.0%
10.	25	89.3%	3	10.7%	16	57.1%	9	32.1%

Because of the disturbing nature of the setting (darkness), however, the behavior of the animals in this instance cannot be regarded as an adequate measurement of the strength of either discriminative ability or gregarious tendency. The tendency which both chicks and ducks exhibited in going to like rather than unlike kind immediately after turning on the lights affords a much better measurement of both their ability to discriminate like and unlike kind and the relative strength of their gregarious tendencies.

Since this differentiated behavior appeared from the first to the tenth day after hatching, against a background of equivalent environmental conditions from birth, it must be regarded as innately differentiated.

The following tables (Tables 4 and 5) show the positional responses for each of the 10 daily series. It will be noted that there is an increased tendency to approach like kind after the light had been turned on. The ducks also display a slightly increasing tendency to approach like kind throughout the 10 daily series.

V. EXPERIMENTAL STUDIES: RESPONSE TO INITIAL PERCEPTION OF WATER

A. Constitutionally Differentiated Behavior of Chicks and Ducks in Their Initial Response to Water

The aim of this experiment is to investigate certain apparently innately differentiated perceptual abilities which are closely integrated with structuro-functional differences. Thus the *perception* of water, just as the *perception* of like and unlike kind, does not depend immediately upon differences in bodily structure or function. Responsive *tendencies* and *abilities*, however (such as swimming) do. At first it might appear somewhat absurd to investigate experimentally the relative behavior of chicks and ducks in response to water. "As a duck takes to water" has become a universally familiar adage. But to grant that "of course a duck will take to water" is to *grant a priori* the innate character of such tendency or predisposition, integrated with structuro-functional organization.

Further, we are not concerned with the fact that their behavior *after entering* water is essentially different from that of chicks under similar conditions, nor yet are we concerned with the duck's apparent greater "enjoyment" of water. The point with which we are here concerned is, how do ducks and chicks respond to water before ever experiencing it, and why? Does this predisposition or tendency to enter water represent an innate aspect of "duck organization," or can it be interpreted adequately in terms of environmental "conditioning?"

Chicks and ducks differ, innately, in their structuro-functional organization, which obviously *adapts* the latter much better than the former to a semi-aquatic mode of living. It seems probable that we should regard native adaptability as closely integrated with *predisposition* toward adaptive behavior. It is further quite probable that predisposition must be thought of in terms of response to *initial perception* or "facility of response" to perceptual experience. Ducklings are usually and normally reared in close contact with older animals from which they may be thought to "learn" certain behavior. If, however, we rear chicks and ducks in an equivalent environment, from birth, their differentiated behavior in response to their initial perception of water may be interpreted in terms of innate predisposition.

The specific problems therefor—given an equivalence of environ-

mental experience, from birth, how will chicks and ducks respond to their *initial* perception of water? Further, with respect to our major criteria for innate tendencies adopted for this investigation as a whole, namely, that of "facility of response" and "ease of learning," with what facility or difference in facility do chicks and ducks respond to water, or "learn" to approach water, especially when encouraged to do so in order to obtain food?

1. *Apparatus*

The apparatus used in this experiment was a "water bath" or tank. It was constructed on a wooden frame, four feet long by two feet wide. This was covered with a screen of quarter-inch wire mesh, to serve as a floor. The purpose of this screen was to avoid the accumulation of food particles and water on the floor, as well as for sanitation. The entire frame was then enclosed with common chick wire, extending one foot above the screen floor (see Figure 10). Twelve inches from one end of this cage a hole was cut in the floor to accommodate a water-pan, 11 x 17 x 5 inches. Wooden

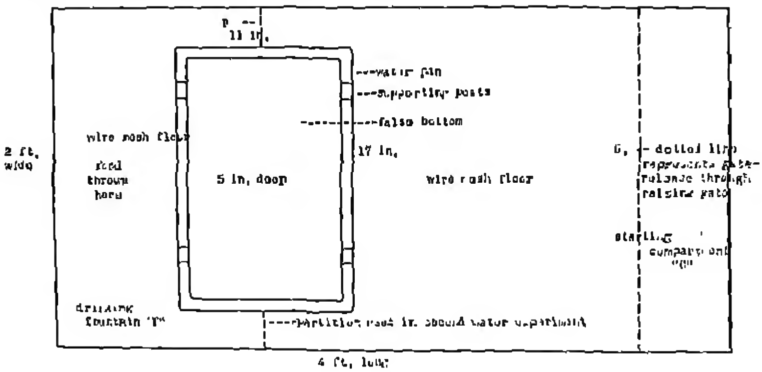

FIGURE 10

WATER BATH: THE RESPONSE OF CHICKS AND DUCKS TO THEIR INITIAL PERCEPTION OF WATER

Experiment VI. Animals pass through or around water to obtain food.

Experiment VII. Animals must pass through water to obtain food (Partition "*p*").

Three ducks and three chicks were released simultaneously from compartment "*C.*"

Behavior was noted with respect to initial response to water.

Behavior was noted with respect to increasing depths of water—food being thrown while the animals were restrained in compartment "*C.*"

supports were inserted to support this pan. This left a space of three and one-half inches at each end of the water pan across which it was possible for the animals to pass without entering the water. A wire gate "*G*" which could be readily lifted, enclosed a six-inch strip at the far end of the cage, which was used as a starting compartment.

A false bottom was then provided for the water-pan by cutting a piece of galvanized iron to a size which would just fit inside the pan. Four wooden strips, one by one-half inches and long enough to reach from the bottom of the pan to just above the surface of the water when the pan was filled, were placed, one near each corner of the pan. They were connected with similar wooden strips across the bottom of the pan for support. Narrow notches were then cut in the vertical strips, into which the edge of the galvanized sheet could be fitted. The whole arrangement was designed so that the false bottom (the galvanized sheet) could be inserted just above or just below the surface of the water, or could be lowered into the water at quarter-inch intervals at will.

Nine chicks and nine ducks, four days old, were used. They had no previous experience with water except in the saucer of their inverted water jar, in which they could just wet their bills.

2. *Procedure*

a. Voluntary approach to the water. The apparatus was placed on the floor (minus the false bottom), levelled, and the pan filled with water, the temperature of the water being brought to room temperature.

The nine chicks and ducks were divided into three groups of three each. For each trial six animals, three of each kind, were placed in compartment "*G*." After a two-minute orientation period they were released by raising gate "*G*." Immediately upon release, all the animals would run to the water-pan together, the ducks slightly in the lead. (We have already noted the greater impulsiveness of the ducks.) In each case, *all the ducks* entered the water without hesitation, some of them swimming, emerging, and re-entering. Two of the nine ducks engaged in repeated vigorous diving. In each case *none of the chicks* entered, though one, coming too near the edge, fell in. Its apparently terrified behavior and frantic efforts to escape were in marked contrast to the behavior of the ducks. (This point is relatively unimportant, as we are con-

cerned with response to initial *perception*, not with behavior after entering, which depends upon structural rather than psychological differences.)

b. Crossing water to obtain food, with choice of going around. The false bottom was adjusted so as to be just level with the floor and slightly above the level of the water (three other chicks and three other ducks, without previous experience with water) were placed in the starting compartment "*C*." Food (standard chick scratch feed), was scattered over a thick covering of paper over the wire floor, across from the starting place. This was done in full view of all the animals.

The animals were then released at three-minute intervals and their behavior noted. With each release the false bottom was lowerd one-quarter inch.

3. *Results*

As the results are typical they will be given in some detail for the first 10 trials.

(*a*) False bottom above water—dry

All walk across for food.

(*b*) False bottom just wet

All walk across without hesitation.

(*c*) ¼ in. water

Ducks walked through.
Chicks—one walked through, one around, another entered water, then jumped rest of way across.

(*d*) ½ in. water

Ducks walked through. Chicks walked around.

(*e*) ¾ in. water

Ducks drink, then enter. One squats in water.
Chicks—all walk around.
Ducks continue to stand in and drink.
One eats while standing in water, and drinks.
All ducks get mouthful of food, then wade and drink.
Chicks do not enter at all.

(*f*) 1 in. water

Ducks immediately enter and walk across.
Chicks go around.
One duck begins to make head and body movements and throw water with head.
Another duck does same 15 sec. later.
Two largest of three ducks remain in water to play.
Ducks do not eat at all except for one or two mouthfuls each—busily engaged in water.
Chicks eat, but do not drink or approach to water.

(*g*) 1¼ in. water

Ducks drink 3 or 4 times, then enter, walking.
One duck crosses water for food.

Two other ducks remain in water, and throw it about.
Chicks walk around.
One chick slipped at edge, one foot goes down to bottom, but is quickly withdrawn.
Ducks apparently enjoying water.
Chicks now peeping and walking around dejectedly; stand near water, apparently wet from all the splashing by the ducks.

(*h*) 1½ in. water

Ducks stop at edge to splash—outside water.
All ducks enter and start splashing.
Then all get out to splash some more from the "shore."
Ducks now give no attention to food at all.
Chicks walk around, peck listlessly at food. Not very hungry apparently, or very contented. Are splashed wet.
Ducks get out and start preening themselves.

(*i*) 1¾ in. water

(Lettuce is thrown on paper, on far side of water.)
Duck 1 *runs* across water for the lettuce.
A chick steps in at edge of water and promptly *jumps* the rest of the way across.
Duck 3 comes for a drink. Does not at once enter. Finally enters.
Chick 3 comes to edge and goes back.
Other chicks remain where they are.
Duck 2 remains where he is and preens self.
Duck 1, the largest of the three, runs back and forth through the water.

(*j*) 2¼ in. water

Duck 1 enters, stands, returns. Walks through. No attempt to swim.
Chicks do not even approach. (Not hungry?)

4. *Summary of Results*

1. With equivalent environmental background, chicks and ducks display distinctly different behavior with respect to entering the water. Not at any time did a duck show hesitation to enter the water. There was rather a distinct tendency throughout to re-enter and disport themselves in the water.

2. With a wet floor, the chicks did not hesitate. But, with water one-quarter of an inch deep the chicks showed immediate and decided aversion to entering, one walking through, one going around, and one jumping across.

3. Not once, after the water was deepened beyond a quarter inch did a chick enter to cross. In every instance the chicks walked around.

4. Not once, on the other hand, did a duck attempt to avoid the water, or to follow the example of the chicks in going around. In every instance they went into or through the water without hesitation, frequently re-entering for considerable periods at a time.

5. After the seventh release, the water now 1¼ in. deep, the

ducks seemed to have "forgotten" all about the food, and disported themselves in the water almost continuously.

6. Throughout the whole experiment, no duck made any attempt to swim, the nearest approach being that of Duck 1, which "squatted" in the water at ¾ in. depth. In the previous experiment, when the ducks voluntarily entered the water for their first experience in it (without the false bottom) they immediately *swam*, with no apparent attempt to *walk* in it.

7. This was, of course, the first experience of any of these animals with water other than that in their drinking fountain. In this experiment the chicks confined their drinking almost entirely to the fountain; the ducks almost entirely to the water in the pan, with occasional drinks from the fountain.

8. Throughout this experiment the ducks again showed their greater impulsive nature in that they appeared more eager than the chicks for release toward the water, and in that they were usually the first to reach the water.

5. *Conclusion*

The conclusion drawn from this experiment are incorporated in the conclusions of the next experiment, for the water experiment as a whole.

B. Experiment V: Constitutionally Differentiated Behavior of Chicks and Ducks in Learning to Cross Water for Food

This experiment is in every way similar to the foregoing one except that in this instance the animals are prevented from going around the water to obtain food. If they are to eat they *must* cross the water.

1. *The Problem*

Given an equivalent environment from birth, and "equal opportunity to learn," how will chicks and ducks differ in *facility of response*, (*a*) to their initial perception of water, and (*b*) to water which must be crossed for food? What will be the effect when the animals are thus "encouraged" to cross the water?

The behavior of the ducks may be anticipated from the results of the foregoing experiment. The problem, therefore, centers primarily around the chicks. With not only "equal opportunity" to learn, but

even with "forced opportunity" to learn, any differentiation in behavior tendency must be rooted in innate constitution rather than in environmental experience.

2. *Apparatus*

The apparatus used in this experiment (Figure 10) is the same as that in the previous water experiment, except that a wire partition, one foot high, and extending from the side of the cage to the water pan at "*p*" prevents the animals from going around the water. The animals must thus cross the water if they are to obtain food.

Eighteen animals were used (not used in any other experiment), nine chicks and nine ducks, in groups of three each. Experiments were begun with these animals at 12 days of age. None of them had ever experienced any other water than that of their drinking fountain—an inverted pint fruit jar in a shallow saucer.

3. *Procedure*

Each group was placed in the apparatus and observed three times daily, for five days, at their regular feeding time, 8:00 A.M., 12 M. and 4:00 P.M. With each feeding for the three groups the false bottom in the water pan was lowered ¼ inch. The first time the groups were entered, the false bottom was adjusted so as to be just above the water level, and dry. With the next feeding, it was lowered so as to be just wet. With each succeeding feeding the false bottom was lowered ¼ inch more. This increase of depth three times daily was in order to reach "swimming depth" without too long a period of growth or development intervening.

4. *Results*

The results are very similar to those in the previous water experiment. The ducks enter the water readily while the chicks attempt to avoid it. When the water has reached a depth of ¼ inch they attempt to go around it. Prevented by the barrier they do go through the water, usually entering after some hesitation and often with little jumps. When the water has become still deeper they attempt to jump or fly across, often from the top of the post near the barrier. Frequently they put one foot tentatively in the water and quickly withdraw it before finally hurrying across. The total behavior, recorded in some detail (but omitted here), indicates a great difference in behavior of chicks and ducks.

a. Progressive development of behavior tendencies. As the experiment progressed the ducks displayed a rapidly increasing repertoire of bodily movements, characteristic of ducks behavior. At first they merely walked in the water, stopping for frequent drinks and straining food particles from the surface of the water with their bills. As the water deepened they spent much time throwing water with quick jerks of the head. At times they would stand on one foot and scratch the face with the other for considerable lengths of time. Apparently they enjoyed the water most at a depth where they could scratch, throw water, and preen while still being able to stand firmly on the bottom.

The ducks spend more and more time in the water with each series. The chicks, however, appear more and more reluctant to even approach the water. When they acquired the method of hop-skipping or jumping across they were less hesitant in approaching though their reluctance to enter was no whit lessened.

The transition from walking to swimming movements (ducks) seemed to be very gradual, under the conditions of gradually increasing depth. As the water became deeper and offered more support to the body, the ducks continued to walk with the feet less and less firmly on the bottom of the pool. As it became deeper still only the toes touched the bottom. Increased floating and decreased walking seemed to lead gradually to more swimming-like movements of the feet until finally the ducks were swimming, apparently with little notice of the change, except for gradually holding feet and legs in a slightly different position, and for greater mobility.

Chicks can swim, after a fashion, though with extreme reluctance, to say the least. One chick, fell accidentally into the water. Its frantic behavior was in sharp contrast to that of the ducks in the water. It moved back and forth along the edge of the pool for fully a half-minute. Fatigued, and unable to escape, it sank deeper and deeper. Upon being rescued it stood in a dazed manner, shook its head with vigorous lateral movements at five- to ten-second intervals for nearly an hour, when the shaking gradually diminished, and the chick recovered. Some ducks, of equal age, upon being *placed* in deep water for the first time, also showed considerable excitement, though quite different from the frenzy of the chicks. Obviously structuro-functional factors, such as buoyancy, resistance to cold, etc., are important aspects of the total organization.

The greater impulsiveness of the ducks is shown in their initial

and subsequent response to the perception of water and in their headlong entry. It recalls their impulsive "launching forth" from the little boxes when they were first removed from the incubator, as noted in the preliminary observations.

b. Axial head-tail movements. One of the most interesting phenomena of duck behavior is the lateral shaking of the head and the vertical movements of the head and neck.

When disporting themselves in shallow water the ducks frequently made twisting vertical movements of the head and neck, a movement used in throwing water over the body. This movement continued as a quivering of the body, proceeding axially, *from head to tail,* and ending in a vigorous lateral shaking of the tail.

Upon *emerging from the water,* however, a vigorous lateral wiggling of the tail *precedes* a vigorous "ducking" movement of the head, the movement being characterized by an obvious quivering of the body, and proceeding axially *from tail to head.*

In observing the behavior of the ducks in the earlier "choice" experiments, it was noted that egress from the tunnel was almost invariably followed by a brief pause in passage, during which the tail was wiggled laterally in very vigorous fashion, the quivering proceeding axially throughout the body *from tail to head,* ending in a vigorous vertical "ducking" of head and neck.

Half grown and adult ducks make similar movements upon hearing running water or upon seeing a pool of water though restrained from entering. In these instances the quivering proceeds *from head to tail.* The behavior of adult ducks in courtship is characterized by similar movements, in the same direction (as noted in the preliminary observations).

Observations seem to indicate that *head-to-tail* movements have a *forward reference,* towards an action as yet incomplete. Tail-to-head movements seem to have a *backward reference,* following the completion of an act.

5. *Conclusion*

The water experiment affords striking evidence of innately differentiated behavior of chicks and ducks.

1. They differ in their response to the *initial perception* of water; the ducks entering immediately, without previous experience; the chicks consistently refraining from entering upon their *initial* perception, even in the face of the example of the ducks, whom they accompanied to the water's edge.

To be sure, the ducks are more impulsive in their behavior. But even though their behavior should be regarded in this light, we nevertheless have an aspect of differentiated behavior of chicks and ducks. Other observations, however, including body posture and the differentiated behavior of chicks and ducks in the "dust-mud" situation in one of our exploratory observations, indicate that perceptual discrimination plays an important part in their behavior.

2. They differ in their *ability to "learn"* to enter water not only with "equal opportunity to learn," but with conditions weighted in favor of chick learning, i.e., encouraging them to enter the water by preventing their going around, and by further encouraging them to cross the water in order to obtain food (the ducks entered readily without either of these incentives), chicks and ducks differ markedly in the *facility with which they learn to approach* water. Further, this difference is indicated in their *initial* approach to water, and is not dependent upon the experience of having entered.

3. These two criteria, that of "appearance without practice," and that of relative "ease of learning," therefore combine to differentiate unequivocally the innate behavior tendencies of chicks and ducks.

4. Chicks apparently have much less sensitivity in their feet than do the ducks. They require a stronger electric shock in their feet, and the exploratory observations showed them to be much less sensitive to the water they inadvertently stepped in. Yet, in strong contrast to the ducks, they showed a strong disinclination to enter water.

5. As these differences in behavior tendencies cannot be due to environmental differences or experience, either before or during the experiment, they must be regarded as being rooted in differences in innate constitution or the "nature" of chicks and ducks.

6. This difference in initial response to water, and in the relative "ease of learning" to enter, in the face of equivalent environmental experience from birth, affords striking evidence for viewing structure, function, predisposition, and discriminative or cognitive ability as integral aspects of innate organization differentiating animal behavior.

VI. GENERAL CONCLUSIONS

These experiments were conducted with a view to testing McDougall's theory that certain perceptual abilities, affective attitudes, and behavior tendencies, as aspects of total behavior, have their basis in innate constitution. Our experimental observations indicate that this viewpoint is substantially sound. Differences in perceptual abilities are, of course, not directly observable, but can only be inferred from differentiated behavior. For this reason we have attempted throughout all the experiments to place emphasis upon behavior tendencies from which differences in *perceptual ability* might possibly be inferred, rather than those involving gross anatomical and functional differences.

We have not attempted to distinguish between "learned" and "unlearned" behavior *as such*, both because of the difficulty of making such a distinction and because we view behavior in the light of an integrated, progressive development, birth being but an arbitrary point in the process. The *facility* with which certain tendencies and abilities appear, in the face of equivalent environmental experience is therefore regarded as a more suitable criterion of innateness than the unaided, elusive one of "unlearnedness."

Interested primarily in the innate psychological basis of *human* individual and social life, we found it desirable to use animals because of the greater ease of manipulation and control of the variable factors involved, i.e., heredity and environment. It seems reasonable to suppose that any conclusions drawn with respect to the origin and development of psychological traits are applicable in varying degrees to both human and animal psychology. Further, we are confident that if it can be demonstrated that certain general or specific behavior tendencies, are rooted in innate constitution, these factors must be taken into account in all educational and social theory and practice. For example, as long ago as 1914, Hahn and Thorndike (27) concluded that, "Nature, not nurture, seems to be the chief cause of the differences in ability to add found in children of the same school grade. Equalizing opportunity does not seem to equalize achievement."

Our various observations permit us to draw the following conclusions:

1. Chicks and ducks, reared under equivalent environmental conditions, from birth, display different tendencies in their behavior.

They tend more readily to approach their own kind, they differ in the strength of this gregarious tendency, and they respond differently to their initial perception of water.

2. From these directly observed behavioral differences, especially with respect to like kind, we infer a necessary innate difference in discriminative or cognitive ability.

3. As this behavior is differentiated through *facility of response* or "ease of learning" in the face of equivalent environmental conditions from birth, these differences in behavior tendencies must be regarded as rooted in differences in innate constitution rather than in differences in environmental "conditions," though the environment, as an aspect of the total situation, does, of course, exert a very profound modifying influence.

4. There seems to be direct evidence that these tendencies and abilities, observed and inferred, respectively, from differentiated behavior, represent a process of maturation and progressive development. There is therefore no good reason to suppose that all differentiated behavior must necessarily be present, full-blown, immediately at birth, in order to be regarded as rooted in innate constitution. Post-natal behavior tendencies must be regarded as aspects of a continuing, progressive physiological and psychological development, which has its roots in pre-natal organization. Accepting "birth" merely as a more or less "natural" yet arbitrary point in this progressive development, we find that chicks and ducks do give evidence of differentiated behavior tendencies and discriminatory abilities immediately after birth.

5. We conclude, finally, that both human and animal behavior represent a developmental process, growing out of the matrix of pre-natal organization, in which bodily structure and function, adaptability, predisposition, and perceptual ability are integrally related. And we will do well in our educational practice to recognize that development is a continuing process in which all behavior must be regarded as the integrated expression of innate factors and environmental experience.

REFERENCES

1. Adams, D. K. Studies of adaptive behavior of cats. *Comp. Psychol. Monog.*, 1929, **6**, No. 1. Pp. 168.
2. Allport, F. H. Behavior and experiment in social psychology. *J. Abn. Psychol.*, 1919, **14**, 297-306.
3. ———. Social Psychology. New York: Houghton Mifflin, 1924.
4. Allport, G. W. Personality, A Psychological Interpretation. New York: Holt, 1937.
5. Allport, G. W., & Schanck, R. L. Are attitudes biological or cultural? *Charac. & Personal.*, 1936, **5**, 195-205.
6. Ayers, E. E. Instinct and capacity. *J. Philos.*, 1921, **17**, 561-566.
7. Bernard, L. L. The misuse of instinct in the social sciences. *Psychol. Rev.*, 1921, **28**, 96-119.
8. ———. An Introduction to Social Psychology. New York: Holt, 1926.
9. Berry, C. S. An experimental study of imitation in cats. *J. Comp. Neurol. & Psychol.*, 1908, **18**, 1-25.
10. Bird, C. The effects of maturation upon the pecking instinct of chicks. *Ped. Sem.*, 1926, **33**, 212-234.
11. Breed, F. S. The development of certain instincts and habits of chicks. *Behav. Monog.*, 1911, **1**, No. 1. Pp. 78.
12. Conradi, E. Song and call notes in English sparrows when reared by canaries. *Amer. J. Psychol.*, 1905, **16**, 190-199.
13. Craig, W. Male doves reared in isolation. *J. Anim. Behav.*, 1914, **4**, 121-133.
14. Craig, E. E. Appetites and aversions as constituents of instincts. *Biol. Bull.*, 1918, **34**, 91-127.
15. Darwin, C. A. A posthumous essay on instinct. In Romanes, G., *Mental Evolution in Animals.* New York: Appleton, 1884.
16. ———. Origin of the Species. New York: Collier, 1909.
17. Dashiell, J. F. Fundamentals of Objective Psychology. New York: Houghton Mifflin, 1928.
18. Drury, M. B. Can Gestalt theory save instinct? *J. Gen. Psychol.*, 1931, **5**, 88-94.
19. Dunlap, K. Are there any instincts? *J. Abn. Psychol.*, 1919-1920, **14**, 307-311.
20. ———. The identity of instinct and habit. *J. Philos.*, 1922, **19**, 85-94.
21. Ellwood, C. A. The Psychology of Human Society. New York: Appleton, 1925.
22. Fabre, J. H. The Mason Bees. (Tr. by Teixeira de Mattos.) New York: Dodd, Mead, 1914.
23. Faris, E. Are instincts data or hypotheses? *Amer. J. Sociol.*, 1921, **27**, 184-196.
24. Fiske, J. The Meaning of Infancy. Boston: Houghton, 1911.
25. Gault, R. H. Social Psychology: The Basis of Behavior Called Social. New York: Holt, 1923.

26. GEIGER, J. R. Must we give up instincts in psychology? *J. Philos.*, 1922, **19**, 94-98.

27. HAHN, H. H., & THORNDIKE, E. L. Some results of practice in additions under school conditions. *J. Educ. Psychol.*, 1914, **5**, 65-84.

28. HINGSTON, R. W. G. Problems of Instinct and Intelligence. London: Macmillan, 1929.

29. HOBHOUSE, L. T. Mind in Evolution. London: Macmillan, 1926.

30. HOCKING, W. E. The dilemma in the concept of instincts. *J. Abn. Psychol.*, 1921, **16**, 73-96.

31. HOLT, E. H. Animal Drive and the Learning Process. New York: Holt, 1931.

32. HOWELLS, T. H. Heredity as a differential element in behavior. *Univ. Colorado Stud.*, 1933, **20**, Nos. 2 & 3.

33. HOWELLS, T. H., & VINE, D. O. The innate differential in social learning. *J. Abn. & Soc. Psychol.*, 1940, **35**, 537-548.

34. HRDLICKA, A. Some reflections regarding human heredity. *Proc. Amer. Phil. Soc.*, 1935, **5**, 298-312.

35. HUNTER, W. S. The modification of instincts from the standpoint of social psychology. *Psychol. Rev.*, 1920, **27**, 247-269.

36. ———. The modification of instincts. *J. Philos.*, 1922, **19**, 98-101.

37. JAMES, W. Principles of Psychology. New York: Holt, 1890.

38. KANTOR, J. R. The problem of instincts and its relation to social psychology. *J. Abn. & Soc. Psychol.*, 1923, **18**, 645-664.

39. KATZ, D. Some fundamental laws of psychology of needs: Hunger. *Charac. & Personal.*, 1935, **3**, 312-326.

40. KING, L. S. Instinct as an explanatory concept. *J. Abn. & Soc. Psychol.*, 1934, **29**, 172-180.

41. KROUT, M. Heredity and environment and developmental process. *Psychol. Rev.*, 1931, **38**, 187-211.

42. KRUGER, M. S. A contribution to the theory of instinct. *Brit. J. Med. Psychol.*, 1932, **12**, 283-294.

43. KUO, Z. Y. Giving up instincts in psychology. *J. Philos.*, 1922, **18**, 645-664.

44. ———. How are our instincts acquired? *Psychol. Rev.*, 1923, **29**, 344-365.

45. ———. A psychology without heredity: The net result of the anti-heredity movement in psychology. *Psychol. Rev.*, 1929, **36**, 181-199.

46. ———. The fundamental error of the concept of purpose and the trial and error fallacy. *Psychol. Rev.*, 1928, **35**, 414-433.

47. ———. Genesis of cat's response to the rat. *J. Comp. Psychol.*, 1930-31, **11**, 1-35.

48. ———. The ontogeny of embryonic behavior in aves: I. The chronological and general, natural behavior of the chick embryo. *J. Exper. Zoöl.*, 1932, **61**, 395-430.

49. ———. The ontogeny of embryonic behavior in aves: II. The mechanical factors in the various stages leading to hatching. *J. Exper. Zoöl.*, 1932, **62**, 453-487.

50. ———. The ontogeny of embryonic behavior in aves: III. The structural and environmental factors in embryonic behavior. *J. Comp. Psychol.*, 1932, **13**, 245-271.

51. ———. The ontogeny of embryonic behavior in aves: IV. The influence of embryonic behavior upon the behavior after hatching. *J. Comp. Psychol.*, 1932, **14**, 109-122.

52. ———. The ontogeny of embryonic behavior in aves: V. The reflex concept—in the light of embryonic behavior in birds. *Psychol. Rev.*, 1932, **39**, 499-515.

53. LEWES, G. H. Problems of Life and Mind. (Author give city & publisher.)

54. LEWIN, K. Environmental forces in child behavior and development. In Murchison, C. (*Ed.*) *Handbook of Child Psychology*. Worcester, Mass.: Clark Univ. Press, 1931. (Pp. 94-127.)

55. LILLIE, F. R. The Development of the Chick. New York: Holt, 1919.

56. LUNDHOLM, H. *Conation and Our Conscious Life.* Durham: N. C.: Duke Univ. Press, 1934.

57. MACE, C. A. Faculties and instincts. *Mind*, 1931, **40**, 37-48.

58. MCCORMICK, T. C. A point of view on instincts in social psychology. *J. Abn. & Soc. Psychol.*, 1931, **26**, 102-105.

59. MCDOUGALL, W. The nature of instinct. *Psychol. Bull.*, 1923, **20**, 200-218.

60. ———. The use and abuse of instinct in social psychology. *J. Abn. & Soc. Psychol.*, 1922, **16**, 285-333.

61. ———. Outline of Psychology. New York: Scribner, 1923.

62. ———. Dynamic principles of gestalt psychology. *Charac. & Personal.*, 1936, **4**, 232-244, 319, 334; **5**, 60-82, 131-148.

63. ———. Social Psychology. Boston: Luce, 1926.

64. ———. The Appolonian and the Dionysian Theories of Man. In *Religion and the Sciences of Life.* Durham, N. C.: Duke Univ. Press, 1934.

65. METFESSEL, M. Objective studies of roller canary song. *Psychol. Bull.*, 1935, **32**, 716-717.

66. MOSELEY, D. The accuracy of the pecking responses in chicks. *J. Comp. Psychol.*, 1925, **5**, 75-97.

67. MUENZINGER, K. F. The primary factors in learning. *Psychol. Rev.*, 1931, **38**, 347-358.

68. ———. Psychology, the Science of Behavior. Denver, Colo.: World Press, 1940.

69. MURCHISON, C., *Ed.* Handbook of Social Psychology. Worcester, Mass.: Clark Univ. Press, 1935.

70. MURPHY, G. Experimental Social Psychology. New York: Harper, 1934.

71. OGBURN, W. F. Discussion of Prof. Kantor's and Prof. McDougall's papers. *J. Abn. & Soc. Psychol.*, 1933, **28**, 50-77.

72. PECKHAM, G. W., & PECKHAM, E. G. Wasps, Social and Solitary. Boston: Houghton Mifflin, 1905.

73. RICHTER, C. P. Some observations on the self-stimulating habits of young wild animals. *Arch. Neurol. & Psychiat.*, 1925, **13**, 724-729.

74. ROMANES, G. J. Mental Evolution in Animals. With a posthumous Essay on Instinct by Charles Darwin. New York: Appleton, 1884.

75. ———. Instinct and acquisition. New York: *Nature*, 1875, **12**, 553-554.

76. ———. Animal Intelligence. New York: Appleton, 1888.

77. ROSENOW, C. One more definition of heredity and instinct. *Psychol. Rev.*, 1928, **35**, 434-439.

78. SCHJELDERUP-EBBE, T. Social behavior in birds. In Murchison, C. (*Ed.*) *Handbook of Social Psychology*. Worcester, Mass.: Clark Univ. Press, 1935.

79. SEWARD, G. H., & SEWARD, J. P. Internal and external determinants or drives. *Psychol. Rev.*, 1937, **44**, 349-363.

80. SHEPARD, J. F., & BREED, F. S. Maturation and use in the development of instincts. *J. Anim. Behav.*, 1913, **3**, 274-285.

81. SMITH, S., & GUTHRIE, E. R. General Psychology in Terms of Behavior. New York: Appleton, 1921.

82. SPALDING, D. A. McMillan's Mag., Feb., 1873. Quoted in Romanes, G. J., *Mental Evolution in Animals*. New York: Appleton, 1884.

83. ———. Instinct with original observations on young animals. *McMillan's Mag.*, 1873, **27**, 282-293.

84. ———. Instinct and acquisition. *Nature*, 1875, **12**, 507-508.

85. ———. Instinctive antipathy of the cat to the dog. *Nature*, **11**, 507.

86. STONE, C. P. The congenital sexual behavior of the young male albino rat. *J. Comp. Psychol.*, 1922, **2**, 95-153.

87. ———. McDougall's methodological concept of instinct. In *Methods of Soc. Sc.* Chicago: Univ. Chicago Press, 1931.

88. THORPE, L. P. Psychological Foundations of Personality. New York: McGraw-Hill, 1938.

89. TOLMAN, E. C. The nature of instinct. *Psychol. Rev.*, 1923, **20**, 200-218.

90. ———. Can instincts be given up on psychology? *J. Abn. & Soc. Psychol.*, 1922, **17**, 139-152.

91. ———. Instinct and purpose. *Psychol. Rev.*, 1920, **27**, 217-233.

92. ———. Purposive Behavior in Animals and Men. New York: Century, 1932.

93. THORNDIKE, E. L. Elementary Psychology. New York: Columbia Univ., 1913.

94. UHRBROCK, R. S. Popular usage of the terms "instinct" and "instinctive." *Sci. Mo.*, 1932, **34**, 545-546.

95. WATSON, J. B. Psychology from the Standpoint of a Behaviorist. Philadelphia: Lippincott, 1919.

96. WELLS, W. R. The value for social psychology of the concept of instinct. *J. Abn. Psychol.*, 1922, **16**, 334-343.

97. ———. The meaning of "inherited" and "acquired," in references to instinct. *J. Abn. Psychol.*, 1922, **17**, 153-161.

98. ———. The anti-instinct fallacy. *Psychol. Rev.*, 1923, **30**, 228-234.

99. WHEELER, W. M. Ants. New York: Columbia Univ. Press, 1923.

100. WHITMAN, C. O. The behavior of pigeons. Carnegie Institution of Washington: 1919, Publication No. 257.

101. ———. Genetic Evolution in Pigeons: The Behavior of Pigeons. Carnegie Institution of Washington: 1919, Publication No. 257.

102. WILM, E. C. The Theories of Instinct. New Haven, Conn.: Yale Univ., 1925.

103. WOODGER, J. H. Biological Principles. New York: Harcourt, 1929.

104. WOODWORTH, R. S. Psychology, a Study of Mental Life. New York: Holt, 1921.

105. WYNDHAM, H. S. Contributions to the problem of instinct in text-books of educational psychology. *J. Genet. Psychol.*, 1934, **45**, 238-254.

106. YERKES, R. M., & BLOOMFIELD, D. Do kittens instinctively kill mice? *Psychol. Bull.*, 1910, **7**, 253-263.

Published as a separate and in *Genetic Psychology Monographs*, 1942, 26, 289-392.

AN INVESTIGATION OF THE INTELLIGIBILITY OF THE SPEECH OF THE DEAF*[1]

C. V. HUDGINS

The Division of Experimental Phonetics, Clarence W. Barron Research Department, The Clarke School for the Deaf, Northampton, Massachusetts

F. C. NUMBERS

The Western Pennsylvania School for the Deaf, Edgewood, Pittsburgh, Pennsylvania

*Received in the Editorial Office on January 17, 1942.

[1]The analysis of that portion of the material of this study covering the articulatory errors of Clarke School pupils was presented by F. C. Numbers to the Faculty of the Department of Education, Massachusetts State College, in partial fulfillment of the requirements for the degree of Master of Science.

TS

It would be quite impossible to conduct an investigation of the type described in this paper without the assistance and complete co-operation of a considerable number of people. It is a pleasure, therefore, to acknowledge the high degree of coöperation and sympathetic interest given the authors by the administrative staffs of the Mt. Airy and Clarke Schools. We also are profoundly indebted to the large groups of teachers of both schools, and to the student teachers of the Clarke School who gave liberally of their spare time for the exacting task of auditing the speech records. Mrs. Ruth B. Hudgins rendered valuable assistance with the statistical analyses of the data and in making the drawings.

C. V. Hudgins
F. C. Numbers

FOREWORD

The authors of this treatise have evaluated quantitatively the intelligibility of the speech of 192 deaf pupils from 8 to 20 years of age in two schools for the deaf in which the oral method is used exclusively as a means of classroom instruction. Short, simple sentences, each one comprising a single idea well within the ken of the subjects and requiring several seconds of articulatory effort, were used; thus affording means of appraising other speech characteristics such as phrasing, accentuation, and rhythm, in addition to recording the precision of articulation. All types of errors contributing to unintelligibility have been reduced to two general categories—errors of articulation and errors of rhythm. The authors contend that rhythmical or non-rhythmical utterances affect speech as adversely, if not more so than, errors of articulation per se; and that the "Elements Method," which is used in the schools under consideration, may be responsible in large measure for the sizeable number of errors, primarily in rhythm, accentuation, and phrasing. The authors suggest that in the light of the results yielded "current methods of speech teaching in schools for the deaf" should receive careful consideration as to their effectiveness in the production of intelligible speech.

Whether sub-standard proficiency in the speech of the deaf child can be ascribed mainly to a method of instruction as such may be questioned, due to the difficulty of isolating or controlling all variables in an investigation of such highly complex human behavior. Speech is the most intricately integrated of all human performances in spite of the fact that it is acquired without any formal instruction by the hearing. The congenitally deaf acquire it only through consciously controlled processes and inevitable formal instruction. Undoubtedly, after more definite and explicit knowledge of speech production and the functioning of the speech mechanism in its parts and as an integrated whole are available it should be possible to get more satisfactory speech or so-called better speech from the deaf child. All teachers of deaf children using the oral method know that much is to be desired in their speech. We realize it and are constantly striving for better speech. In accord with this objective procedures

in our school will be most carefully examined in line with the suggestions of the authors and such modifications and alterations adopted as subsequent experience justifies.

THE TRUSTEES OF THE CLARKE SCHOOL FOR THE DEAF

Northampton, Massachusetts

I. INTRODUCTION

This paper presents a detailed analysis of speech samples obtained from 192 deaf pupils between the ages of 8 and 20 years, with hearing losses ranging from a slight impairment to profound deafness. The data were obtained from the pupils of two residential oral schools for the deaf. The purpose of the investigation was two-fold: (*a*) to identify speech errors, to classify them into general categories, and to determine their frequencies; and (*b*) to determine the relative effects of each type of error upon the intelligibility of the speech samples.

The teaching of speech to deaf children should have for its chief purpose the development of an intelligible medium of communication. If this purpose is to be achieved it is essential that the characteristic errors in the speech of deaf pupils, those errors which by their frequency of occurrence and general characteristics offer the greatest hindrances to speech intelligibility, be well understood.

The results of this investigation are of interest primarily to teachers and administrators in schools for the deaf. The materials to be presented, however, should be of interest also to phoneticians, phonologists, speech pathologists, and psychologists, all of whom are interested in one or more aspects of the complex problem of speech and language development.

The speech of deaf and partially deaf pupils exhibits wide variations in degree of intelligibility. Factors upon which intelligibility depends are both general and specific: General factors include speech rhythm, which involves accentuation, grouping and phrasing; voice quality; and intonation. Specific factors include: (*a*) the relative accuracy, in both quality and function, of individual phonemes which combine to form words and phrases; (*b*) the manner in which these phonemes combine or fuse in syllables; and (*c*) the accuracy of coordination between the articulatory organs which produce the phonemes and the chest-abdominal muscles which control the movement of the air-column in the production of syllables.

The teaching of speech to deaf pupils is essentially a matter of teaching a motor skill. The normal child develops this skill early and with little or no specific training since the sounds produced by the speech movements themselves provide ample sensory cues. Without the benefit of the sense of hearing to guide him the deaf child must rely upon visual, tactile and kinaesthetic cues for the control

of speech movements. The speech mechanism is made up of groups of widely distributed muscles which become coördinated into close functional relationship in the process of speech production. The accuracy of speech, or its intelligibility, therefore, depends not only upon the accuracy of each component movement, but even to a greater degree upon the accuracy of timing and synergic action of individual components of the speech mechanism as a whole.

Indeed, the accuracy of individual movements within the speech coördination may vary widely without materially affecting speech intelligibility. A study of the degree of toleration of speech inaccuracies permitted by the normal ear would be most revealing. The types of inaccuracies which occur in the speech of deaf children differ from those which constantly occur in the daily speech of hearing persons. For this reason normal hearing persons who tolerate wide variations of degree of accuracy in the careless speech of their friends have considerable difficulty in understanding the speech of deaf children. With intelligibility the primary goal in teaching speech to deaf children, therefore, it seems essential to determine not only the types of errors which occur but in addition the relative effects of these errors upon speech intelligibility.

II. SOME PREVIOUS STUDIES IN THE SPEECH OF THE DEAF

There has been no previous attempt to the writers' knowledge to work out experimentally the relationship between errors of articulation and rhythm and intelligibility in the speech of deaf children. Comprehensive studies dealing with various aspects of the speech of deaf children have been presented and comparisons have been made between the speech of deaf and normal hearing subjects. Of these studies those of Hudgins (9), Rawlings (12), Voelker (22), and Scuri (13) may be mentioned.

Hudgins (9), 1934, made kymograph recordings of speech breathing movements and buccal pressure tracings of 62 deaf and 25 normal subjects as they repeated phrases of 4, 5, 7, and 9 syllables in length. The records were measured to determine the relative amount of breath used per phrase, and the rate of syllable utterance. The speech of the deaf group showed the following abnormalities: (*a*) slow and labored speech usually accompanied by high chest pressure with the expenditure of excessive amounts of breath; (*b*) prolonged vowels with consequent distortion; (*c*) abnormalities of rhythm; (*d*) excessive nasality of both vowels and consonants; and (*e*) malfunction of consonants with the consequent addition of adventitious syllables between abutting pairs.

Rawlings (12), 1935, using a method similar to that of Hudgins, recorded the speech breathing movements of deaf and normal subjects as they repeated a short prose paragraph instead of the disconnected phrases. His results were similar to those of Hudgins. His method permitted the comparison of the amount of breath consumed while speaking and while breathing quietly for an equal period of time. He found that normal speakers use approximately the same amount of breath while speaking in a conversational tone as they use in quiet breathing over a similar period. Deaf speakers on the other hand use a great deal more breath while speaking.

Scuri (13), 1935, studied several aspects of speech breathing in a group of 50 deaf subjects between the ages of 8 and 18 years. Kymograph records were made during quiet and deep breathing, during the articulation of vowels, while repeating phrases of varying lengths, during spontaneous speech, and while holding the breath at the end of inspiration and again at expiration. Scuri's results agree with those

of Hudgins and Rawlings mentioned above. In addition he found that congenitally deaf children have poor voices resulting from their inability or failure to completely close the glottis. One of the primary reasons for the breathiness of voices of deaf children, he thinks, is due to their inability or failure to bring the vocal folds into the proper degree of approximation for voice production. This defect also accounts in part for the excessive amount of breath which accompanies the speech of deaf children. The fact that among the pupils studied the younger ones just entering school and those who were of graduating age had less difficulty in controlling the laryngeal muscles indicated to Scuri that speech training interfered with the normal functioning of the laryngeal muscles, and that later development tends to restore this function.

Voelker (22), 1935, compared a group of deaf children with their teachers and with hearing children of the same age in terms of pitch changes during speech, "duration of phonation," and length of pauses. He found that 80 per cent of the deaf children had less average pitch changes than normal speakers. The maximum "phonation duration" for deaf children was more than four times that of normal speakers. Deaf children "took almost four times as long to say a sentence on the average as did the normals." Deaf children "used three times as much phonation to say a sentence on the average as the normals."

III. METHODS AND PROCEDURES

The essential methods of the study consisted of: (*a*) obtaining phonographic recordings of speech samples from a large group of deaf children; (*b*) testing the samples for intelligibility by having groups of hearing persons, serving as *auditors*, listen to reproductions of the records; (*c*) analyzing the records for articulatory and rhythmic errors by repeatedly reproducing them and phonetically transcribing the speech materials; and (*d*) determining the relative effects of speech errors upon speech intelligibility by correlating *speech errors* with *errors of auditors*.

Test samples of the speech of 192 deaf pupils between the ages of 8 and 20 years from two oral schools for the deaf were obtained. The samples were electrically transcribed on either aluminum or acetate discs. There are advantages and disadvantages in this method. The disadvantages of this method of testing for speech intelligibility lie in the inherent defects of the electrical recording method, namely, the lack of complete fidelity of recording and reproduction. The loss of accuracy incurred by using this method, however, was minimized as far as possible by using the best type of commercial recording apparatus available at the time of the study, 1936-37. Furthermore, the loss in intelligibility, or accuracy, incurred should be a constant factor operating throughout the entire group of subjects rather than penalizing single individuals. There are two distinct advantages in the method which out-weigh, for our purpose, the disadvantages: (*a*) The auditors listening to the records must get cues from the speech sounds alone without the aid of lip-reading which is an important factor in face-to-face speech tests. (*b*) The speech recordings provide a permanent record of the speech of each individual which can be studied and transcribed at leisure by the experimenter.

A. Test Materials

The test materials used consisted of 1200 simple, unrelated sentences typewritten on cards in sets of 10 and numbered from 1 to 10. The teachers of the pupils to be tested collaborated with the experimenters in constructing the sentences. Difficult and unfamiliar words as well as complicated grammatical constructions were avoided. No attempt was made to "load" the sentences with any particular type of words or sounds. The sentences used were short, simple statements or questions ranging in length from 6 to 12 words. A

sample card chosen at random showing the type of sentences used follows:

1. Sally likes to swim.
2. She bought some bread for her mother.
3. Bob has a little black dog.
4. Polly helps her mother every day.
5. We saw many cows on the farm.
6. The baby bear likes milk.
7. Have you a toy train?
8. Four and five will not make ten.
9. Butter is made from cream.
10. A baby sheep is a lamb.

The sentences used at Clarke were also used at Mt. Airy plus an additional 18 sets.

B. Recording

Each pupil was allowed to read over silently, and at least once aloud, the test sentences before his speech was recorded. Unfamiliar words, or words mispronounced during this preliminary reading were explained or corrected. The cutting head was lifted from the recording disc momentarily between sentences in order to separate each sentence from another by a narrow blank record space and thus make it easier to identify each sentence later. Larger spaces were left on the record between groups of 10 sentences, or the test materials of each subject. All of the recording equipment was set up in semi-soundproof rooms.

The younger children were brought to the room by their teachers who gave them the necessary instructions and assisted generally in directing the test. The name and age of each pupil was recorded on the discs following the test material as a means of indentification. Each record was played back immediately it was made and, if found to be defective, a duplicate was made.

C. Auditing the Records

The records were audited by persons familiar with the speech of the deaf. Teachers in both schools represented by the pupils served as auditors as well as students of the Teacher Education Department, Clarke School, who had been observing deaf pupils for at least several months. From 5 to 10 persons served as auditors for all records;

the average number of auditors was 7. Persons totally naïve with regard to speech of deaf pupils would have been of little value as auditors since they are often distracted by the peculiar voice qualities common to deaf pupils and thus lose much of the content. Special *audition forms* were provided which contained spaces for indentification data and spaces in which each of 10 sentences could be written three times. The auditors received the following instructions:

> *You will hear recordings of the speech of deaf children.* The material consists of short unrelated sentences. Each sentence will be reproduced three times and you will be allowed time between reproductions to write down what you think the child says. Please write down what you *think* the child says after each reproduction, correcting what you formerly wrote if you have reason to believe that you were wrong on your first or second trial. Please do not guess wildly, but write down as nearly as possible what you think the child says. If you do not understand whole words indicate the sounds as nearly as possible.

D. SUBJECTS

One hundred and ninety-two deaf and partially deaf pupils chosen from two oral schools were used in the study. The Clarke School group consisted of 87 pupils between the ages of 8 (one case) and 19 years. The Clarke pupils were evenly divided as to sex, there were 43 girls and 44 boys; these included all of the pupils in the Intermediate and Grammar Divisions; no pupils were tested in the Primary Department. The Mt. Airy groups consisted of 105 pupils between the ages of 8 and 20. Three classes from the upper grades in the Primary Department were included. Of these 43 were girls and 62 were boys. The choice of pupils who were used in the study from Mt. Airy school was left to the principal and to the supervising teachers in each of the three divisions. They were asked to select those who were listed as being either congenitally deaf or those who had become deaf at pre-speech age; those of average intelligence or higher; and those who represented varying degrees of deafness from hard-of-hearing to profound deafness.

The entire group from both schools, therefore, consisted of 192 pupils, 86 girls and 106 boys. Audiograms were available for all pupils studied by which it was possible to group the pupils into three groups according to the degree of hearing loss. The audiometric

classification used by Guilder and Hopkins (6) was used. By this method the pupils were divided into three groups: *A*, *B*, and *C* according to the degree of hearing loss as determined by the audiometer test. The boundaries between groups are arbitrarily fixed by mathematical formulations (6, pp. 73-74). A score of 800, 100 for each of eight octaves (64 to 8192 d. v/sec.) included in the hearing range tested, is considered perfect or normal hearing. Individual scores are based on the sum of the percentage of hearing at each of these eight octaves. The percentage of hearing at any single octave is obtained by dividing the decibel loss by the total intensity range possible at that particular octave and multiplying by 100. By using this system the hard of hearing pupils, Group *A*, are those whose total score ranges from 700 to 426; the partially deaf pupils, Group *B*, have a range between 425 and 226; and the profoundly deaf pupils are those whose total scores range between 225 and 0. In the two school populations there were 21 pupils in Group *A*, the "hard-of-hearing" group; 56 in Group *B*, the "partially deaf" group; 115 in Group *C*, the "profoundly deaf" group. Pupils in Groups *A* and *B* have sufficient residual hearing to benefit by the use of hearing-aids, those of Group *C* may possibly be benefited by the use of hearing-aids but to a much lesser degree than those of Groups *A* and *B*. All of the pupils in Groups *A* and *B* from the Mt. Airy School were using group-hearing aids for all of their classroom work; only a part of these same groups of the Clarke pupils were using hearing-aids at the time of this survey.

E. Analysis of the Data: Methods

The original data consisted of acoustic records of speech samples taken from 192 deaf children. Approximately 1400 *audition-records* were obtained from these by the method of having groups of auditors listen as the acoustic records were reproduced. Each audition-record consists of the auditor's response to 10 sentences reproduced three times.

Speech intelligibility scores were determined from the audition-records as follows: Credit of 10 points was given for each sentence understood by each auditor. No credit was allowed for partially correct auditions. Thus the difficulties involved in attempting to assign objective scores to partially completed sentences were avoided. The three repetitions of each sentence during the process of auditing

the records gave the auditors an oppportunity to correct first impressions; thus the pupil was given a fair chance to obtain a legitimate intelligibility score since he was not penalized when sentences were not understood on the first reproduction. Each auditor, therefore, contributed a single score for each pupil. An average of the total number of auditors' scores was taken as the final intelligibility score.

The maximum score was 100 since each of 10 sentences was given a credit of 10 points.

After the auditions were completed the acoustic records themselves were analyzed by two experimenters who were familiar with methods of phonetic transcription. The method used in this analysis was that of having one of the experimenters reproduce the record while the other transcribed phonetically the sounds spoken by each pupil in his efforts to repeat the sentences. The final transcription was based upon the common agreement of the two experimenters. In the same manner the rhythm of the sentences was transcribed in appropriate symbols just as it was spoken by the pupils. Several reproductions of each sentence were necessary for this type of transcription. One of the two experimenters had copies of the speech material which was being reproduced.

A reliability coefficient of .90 was obtained for this method of determining articulatory errors in the speech tests. The method of determining the reliability was the "split-half" method: The number of errors in the five odd numbered sentences (in sets of 10) were correlated with the total for the five even numbered sentences. A coefficient of correlation of .81, *P.E.* .03, was obtained for the split-half of the sentences. By applying the Spearman-Brown formula (7, p. 419) a reliability coefficient of .90 was obtained for the total number of sentences analyzed.

The phonetic transcriptions of the speech materials contained all of the errors of articulation and errors of rhythm made by the pupils. These data were analyzed for the purpose of relating articulatory errors and *audition errors.* No articulatory error was included in the final data that apparently did not cause at least one auditor to err in his interpretation of the sentence in question. The treatment of the errors of rhythm will be discussed in Section IV, *G.*

F. Consonant Error Categories

The next step in the analysis consisted of analyzing the articula-

TABLE A
KEY ALPHABET[1]

	Consonants		
p	*p*in, cu*p*	zh	azure, rou*g*e
b	*b*in, cu*b*	ch	*ch*oke, ri*ch*
t	*t*en, be*t*	j	*j*oke, ri*dg*e
d	*d*en, be*d*	m	*m*et, hi*m*
k	come, ba*ck*	n	*n*et, thi*n*
g	*g*um, ba*g*	ng	——, thi*ng*
f	*f*an, sa*fe*	l	*l*aid, dea*l*
v	*v*an, sa*ve*	r	*r*aid, dea*r**
t^1h	*th*igh, ba*th*	w	*w*et, ——
t^2h	*th*y, ba*the*	y	*y*et, ——
s	*s*eal, ra*ce*	wh	*wh*en, ——
z	*z*eal, rai*se*	h	*h*am, ——
sh	*sh*ore, ru*sh*		
	Vowels		
o^1o	st*oo*l, y*ou*	ee	b*ee*t, *ea*se
o^2o	w*oo*l, b*oo*k	-e-	l*e*t, *e*dge
o-e	p*o*p*e*, t*o*n*e*	-i-	s*i*t, *i*s
aw	*aw*ed, n*au*ght	-a-	p*a*t, *a*m
a(r)	p*a*rt, *a*lms	a-e	*a*g*e*, p*a*y
-u-	s*u*n, t*o*n	i-e	b*i*t*e*, *ai*sle
ur	*ur*ge, f*ir*st	ou	*ou*t, p*ow*er
-o-	h*o*t, *o*dd	oi	*oi*l, b*oy*

[1]This system is based upon the common English spellings of the several phonemes and familiar to teachers of the deaf as the *Northampton Chart* System (24, pp. 10-11).

**R* is not listed as a final consonant on the Northampton Charts.

tory errors (Table A) and classifying them according to error types or categories. Consonant errors were assigned to seven categories while vowel errors fell into five categories. A brief description of each category follows:

1. Errors involving the confusion of the surd-sonant (voice-breath) distinction. This type of error occurs when *p* is given for *b*, *t* for *d*, *k* for *g*, or *f* for *v*, and *ch* for *j*; in other cases the reverse occurs, thus *b* may be given for *p*, but in general the surd is substituted for the sonant. This type of error may be completely confusing when the error produces a word with a totally different meaning such as *hat* for *had*, *curl* for *girl*, *coat* for *goat*.

2. Substitution of one consonant for another, such as *w* for *r*, *l* for *r*, *t* for *s*; thus *won* is heard instead of *run*.. Any consonant may conceivably be substituted for another but there are definite trends in substitution as will be shown later.

3. Errors involving control of the velum (nasality of consonants

or the lack of it). Thus *m, n,* or *ng* are substituted for *p/b, t/d,* or *k/g*. The reversal of this trend also frequently occurs, thus *p* or *b* is substituted for *m,* or *d* for *n, k* or *g* for *ng*. Any consonant may become nasalized due to the failure of the child to elevate the velum sufficiently to close off the nasal pharynx. Extreme forms of nasality appear in cleft palate speech when the closure of the nasal pharynx is impossible.

4. Errors involving the articulation of compound consonants. These take one of two forms: (*a*) One or more of the members making up the compound may be dropped; thus *street* becomes *st'eet, 'treet* or even *teet, place* becomes *p'ace*. (*b*) The members of the compound are given too much time and spoken too slowly with the result that adventitious syllables are added to the word; thus *snow* becomes *su now, birds* becomes *bir dus, six* becomes *sikus*. This type of error involves both the phonetic structure and the rhythmic pattern of the phrase.

5. Errors involving abutting consonants. In words or phrases where the final or arresting consonant of one syllable abuts with the releasing consonant of the next, e.g. fla*gp*ole, fi*ft*een, deaf children often fail to observe the arresting function of the one consonant and disjoin the articulatory organs before the closure of the following consonant occurs. The result is the addition of an adventitious syllable. Instead of *flagpole* or *fifteen* we hear flag*u*pole, fif*u*teen. Both the phonetic structures and the rhythmic patterns are changed.

6. Non-function of the arresting (final) consonant. The consonant movement is either dropped out entirely or the movement is incomplete, or again it may be so slow that its dynamic effect upon the preceding vowel is lost. In either case the syllable is not arrested and the vowel trails off slowly. Thus *Paul* becomes *pau' . . .*, *house* becomes *hou'. . . .*

7. Non-function of the releasing (initial) consonant. The consonant movement fails to close or to make the proper juncture with the opposing surface thus preventing a sufficient constriction for the air pressure to produce the consonantal effect upon the syllable. The articulatory movements may be made but they are too slow, or the proper closure is not made. They have no resemblance to the consonant movement; and become passive "oral gestures" lacking in all the dynamic qualities of true consonants. The acoustic effect is that of dropping the consonant.

G. Vowel Error Categories

1. Substitution of one vowel for another, thus *paple* for *for people*, *Jane* for *John*.

2. Errors involving diphthongs. These take one of two forms: (*a*) the diphthong is split making two distinct vowels instead of a fusion of the two components. This is caused by the slowly executed movements of articulation which allow too much time for the vowel and the improper coördination of breathing muscles. An example of this type of error is *bau-ee* for *boy*, *da-ee* for *day*. (*b*) One of the components of the diphthong, usually the final member, is dropped. This is caused by a failure to make the complete articulatory movement, thus *I* becomes *ar*, *found* becomes *fond*.

3. Diphthongization of pure vowels. The slowing down of articulatory movements and a continuation of the voice during the transition from one sound to another often has the effect of making diphthongs out of pure vowels. For example, in the phrase "how do you do" the vowels *oo* are continued while the articulatory movements are slowly moving to form the next consonant and the phrase becomes "*how-ee do-ee you-ee do-*."

4. Neutralization of pure vowels. In this case the oral movements required for producing the vowel are not given full value and the vowel appears to fade out, loses its quality and becomes more like the neutral vowel -*u*-. The syllables are usually shortened and are not given their proper degree of stress.

5. Nasalization of pure vowels. Any vowel may be given too much nasal resonance and the tone takes on the quality of "cleft-palate" speech by the failure of the subject to control the velum during the articulation. It is often difficult to determine whether the consonant movement or the vowel is at fault for the nasality is present in both.

IV. RESULTS

The speech errors fall into two general types: (*a*) errors of articulation including consonant and vowel errors; and (*b*) errors of rhythm. Other defects such as defective voice quality, false intonation, and monotonous speech, while present could not be systematically analyzed into quantitative terms and are not included in the study. Peculiarities of voice quality and intonation are common in the speech of deaf children and, of course, exert some influence upon the intelligibility of their speech. While these factors are not included in the quantitative data they are, at least in part, controlled by using as auditors teachers and student teachers familiar with the speech of deaf children. The use of auditors totally unfamiliar with the speech of deaf children would have given undue emphasis to qualitative aspects of the speech samples. The articulatory errors and errors of rhythm, grouping, and accentuation of syllables could be determined directly from the recorded speech samples. Furthermore, the effects of these errors upon speech intelligibility could be determined by checking them against the audition-records.

The results are presented in the following order: (*a*) General analysis of articulatory errors; (*b*) discussion of individual consonant error categories; (*c*) discussion of individual vowel error categories; (*d*) the relative importance of error categories with reference to speech intelligibility; (*e*) analysis of articulatory errors according to degree of hearing loss; (*f*) articulatory errors and age; (*g*) analysis of errors of rhythm; (*h*) the relative difficulty of individual phonemes as determined by frequency of errors in the speech of the deaf.

A. General Analysis of Articulatory Errors

Clarke pupils (87) made 1,963 consonant errors; this was an average of 23 errors per pupil, and amounts to 22 per cent of all consonants in the test material. Mt. Airy pupils (105) made a total of 2,154 consonant errors, or an average of 20 per pupil and approximately 20 per cent of all consonants in the speech material. Vowel errors were fewer among both groups: the averages were 9 and 8 per pupil for Clarke and Mt. Airy respectively, or 12 and 11 per cent of all vowels attempted.

The distribution of articulatory errors according to individual categories is presented in Table 1. This is an over-all presentation

TABLE 1

THE DISTRIBUTION OF ARTICULATORY ERRORS MADE BY 192 DEAF CHILDREN, ACCORDING TO ERROR CATEGORIES, IN TERMS OF: (*a*) TOTAL NUMBER OF ERRORS MADE; (*b*) ERRORS MADE IN TERMS OF PER CENT OF TOTAL; (*c*) NUMBER OF ERRORS POSSIBLE IN EACH CATEGORY; AND (*d*) ERRORS MADE IN TERMS OF PER CENT OF POSSIBLE ERRORS IN EACH CATEGORY

	Clarke 87 pupils				Mt. Airy 105 pupils			
	(1) No. errors made	(2) % of total	(3) No. errors possible	(4) % of possible errors	(1) No. errors made	(2) % of total	(3) No. errors possible	(4) % of possible errors
Consonant Categories								
1. Surd-Sonant distinction	399	20.3	6,352	6.3	848	40.0	7,700	11.0
2. Substitution	149	7.6	8,800	1.7	174	8.1	10,670	1.6
3. Nasality	225	11.5	8,800	4.0	183	8.3	10,670	1.8
4. Compound consonants	355	18.1	920	38.0	289	13.2	1,103	26.0
5. Abutting consonants	121	6.2	2,000	6.0	85	4.0	2,415	3.5
6. Dropping arresting consonants	266	13.5	3,152	8.5	245	11.1	3,760	6.5
7. Dropping releasing consonants	448	22.8	5,670	8.0	332	15.1	6,850	4.8
Totals	1,963	100.0	8,800	22.2	2,154	100.0	10,670	20.2
Vowel Categories								
1. Substitution	444	55.0	6,860	6.5	516	66.0	8,280	6.2
2. Diphthongs	72	9.0	800	9.0	89	11.4	966	9.2
3. Diphthongizing vowels	76	9.5	5,880	1.3	38	4.8	7,100	.5
4. Neutralization	153	19.1	6,860	2.2	67	8.6	8,280	.8
5. Nasality	56	7.0	6,860	.8	73	9.3	8,280	.9
Totals	801	100.0	6,860	11.6	783	100.0	8,280	10.6

of the articulatory errors from the pupils of both schools presented separately. Columns 1 give the total number of errors in each error category; Columns 2 give these same errors in terms of per cent of total errors in all categories. The data in Columns 1 and 2, however, are obviously not sufficient as an analysis since the total number of chances for error in the several categories is unequal. Columns 3 and 4 are presented, therefore, for further clarification. The total number of consonants and vowels in the entire group of test sentences, which can be considered as possible errors in each category, is presented in Columns 3. These totals, therefore, represent the number of chances for error, or the total number of potential errors, in the test material.[2] Columns 4 present for each school

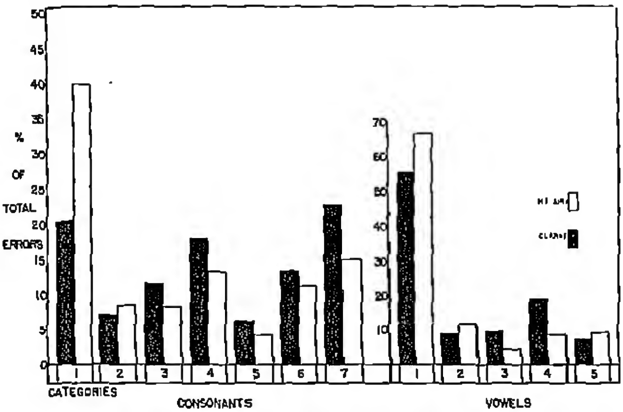

FIGURE 1

CONSONANT AND VOWEL ERRORS AT EACH CATEGORY EXPRESSED IN TERMS OF PER CENTS OF TOTAL ERRORS MADE

Data for the two school groups are shown separately. The height of the columns indicate the frequency of errors in each category in terms of per cent of the total number of errors made.

Categories:

1. Surd-sonant	1. Substitution
2. Substitution	2. Diphthongs
3. Nasality	3. Diphthongization
4. Compound consonants	4. Neutralization
5. Abutting consonants	5. Nasality
6. Arresting consonants	
7. Releasing consonants	

[2]These data were obtained by counting the number of consonants and vowels falling into each category in 100 unselected sentences from the test material.

the errors made in each category in terms of per cent of errors possible.

The data in Table 1 can be presented to better advantage in graphic form. Figure 1 is a graphic presentation of the data of Columns 2 which represent the per cent of total consonant and vowel errors in each category. Consonant error Categories 1 (surd-sonant errors), 7 (releasing consonant errors) and 4 (compound consonant errors) show the greatest per cent of errors in both school groups. The major difference between the two schools appears in Category 1 in which Mt. Airy pupils almost double the errors of the Clarke pupils. Little or no significant differences between the two school groups appear in the other categories.

The vowel errors are similar for the two schools. There is a preponderance of errors in vowel Category 1 (substitution). The other categories, with the exception of 4 (neutralization of vowels) are relatively small and show very little differences.

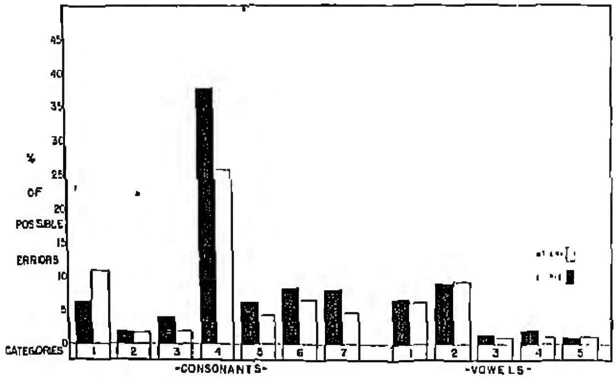

FIGURE 2

CONSONANT AND VOWEL ERRORS IN EACH CATEGORY EXPRESSED IN TERMS OF PER CENT OF POSSIBLE ERRORS

Data for two school groups are presented separately. The height of the columns represents the frequency of errors in terms of per cent of the total possible errors in each category.

Categories:

1. Surd-sonant
2. Substitution
3. Nasality
4. Compound consonants
5. Abutting consonants
6. Arresting consonants
7. Releasing consonants

1. Substitution
2. Diphthongs
3. Diphthongization
4. Neutralization
5. Nasality

In contrast to Figure 1, which presents the errors in terms of per cent of total errors, Figure 2 shows these same errors in terms of per cent of possible errors (Table 1, Col. 4). Two major differences appear when the data is thus presented as compared with Figure 1: In the first place, looking at the error distribution for consonants, Category 4 (errors involving compound consonants) is the most prominent. This means that errors of this type occur far more frequently, in proportion to the chances for such errors, than those of any other category. It means that 38 and 26 per cent (30 per cent for both schools combined) of all compound consonants were malarticulated by the subjects. In the second place, the height of the column representing Category 1 is greatly reduced in Figure 2. Figure 1 shows that approximately 30 per cent of all errors were of the sonant-surd type, but when these data are expressed in terms of per cent of chances for error they are reduced to 6.2 and 11 per cent respectively for the two schools. The problem which is suggested here, that of the relative importance of different types of errors for speech intelligibility, will be discussed later.

The distribution of errors according to per cent of possible errors (Figure 2) also changes the relative height of the columns representing the vowel errors. Category 2 (errors involving diphthongs) now becomes the largest with Category 1 (substitution) second. The remaining categories are relatively small in magnitude.

It is significant to note that in both consonants and vowels, the type of error which is relatively most frequent in terms of possible errors (4 and 2) is the type which involves a fusion of two or more component movements. The important factor in the articulation of compound consonants and compound (diphthong) vowels is the fusion of two or more articulatory movements into a single movement. Components of both members of the compounds are normally present but the rapid articulatory strokes fuse so that the acoustic-phonetic effect is a fusion which blends completely the two or more elements. A slowing up of the component movements tends to separate them. When components of a compound consonant or a diphthong are separated a superfluous vowel is added to the word. Thus the number of syllables in the word is increased, and both the acoustic and the rhythmic patterns of the word are distorted.

It is of interest to note the types of consonants and vowels which make up these individual error categories. Table 2 shows a list of consonants and vowels most frequently malarticulated in each cate-

TABLE 2

LISTS OF 10 CONSONANTS AND VOWELS[1] MOST FREQUENTLY MALARTICULATED IN EACH ERROR CATEGORY RANKED IN ORDER OF HIGHER FREQUENCY

Categories	1		2		3		4		5		6		7	
	*Cl.	*MtA.	Cl.	MtA.	Cl.	MtA.	Cl.	MtA.	Cl.	MtA.	Cl.	MtA.	Cl.	MtA.
							Consonants							
Rank														
1	d	d	r	s	n	t	st	st	n:th	n:th	l	l	h	h
2	b	b	s	r	m	t̄h	pl	fl	n:t	t:k	t	t	l	l
3	g	z	sh	ch	t̄h	m	bl	ny	s:t	t:th	s	d	y	y
4	w	g	ch	sh	f	sh	kl	fr	v:m	k:t	d	z	r	r
5	z	w	j	j	p	d	gr	pl	d:y	l:v	z	s	t̄h	g
6	k	s	b	t̄h	t	s	ny	kl	k:t	s:l	g	n	s	t̄h
7	s	b	d	–ng	s	v	tr	–nt	d:n	n:s	k	t̄h	sh	wh
8	r	j	–ng	v	b	p	fr	sl	f:t	t:m	v	g	f	k
9	p	t	l	y	d	b	br	–ld	n:s	st:m	sh	p	p	f
10	t	wh	g	k	z	g	–nd	gr	t:k	n:d	n	k	wh	w
							Vowels							
Rank														
1	a-e	–i–	i-e	i-e	o^2o	o^1o	–i–	–i–	aw	–a–				
2	i-e	–e–	oi	ou	ee	aw	aw	i-e	–i–	i-e				
3	ee	ee	ou	oi	aw	o-e	o-e	ur	–u–	–u–				
4	–i–	–o–	a-e	a-e	–a–	ur	o^2o	o^1o	ee	–o–				
5	–e–	i-e	—	—	o-e	–u–	a-e	aw	I o^1o	i				

[1]The system of transcription here used is that of the Northampton Charts. (See p. 504).
*Cl. equals Clarke.
*MtA. equals Mt. Airy.

gory presented in rank order. The data are separated for the two school groups. There is a surprising agreement between the two schools in the consonant frequency lists; vowels show less agreement.

B. Analysis of Individual Consonant Error Categories

A discussion of the individual error categories, along with illustrations of the effects of these errors upon auditors taken from audition-records, should be of interest.

1. *Errors Involving Surd-Sonant Consonants*

There are decided acoustic and articulatory differences between the surd and the sonant. During the occlusion of a sonant consonant there is an audible laryngeal tone which distinguishes it from both the complete silence of surd stops and from the high frequency hissing of the surd continuants. In the articulation of sonants air must be moved upward through the glottis during the consonant occlusion even when the mouth is completely closed as in the stops *b*, *d*, and *g*. Normal speakers learn to do this by momentarily depressing the larynx, thereby enlarging the supra-glottal cavity and reducing the buccal pressure. An experimental analysis of the surd-sonant mechanism has been presented by Hudgins and Stetson, 1935 (11). Deaf children find it difficult to control the thoracic and buccal pressures in the manner required for the proper articulation of sonant stops. They can be taught the surd-sonant distinction by special methods (10). The failure to distinguish between these two types of consonants constitutes one of the largest categories of error in the speech of the pupils studied in this survey. The sonant stops *b*, *d*, and *g* are among the four most frequently malarticulated (Table 2, Category 1).

A few illustrations will show the effects of such errors upon speech intelligibility. Since there was no continuity of meaning between the test sentences the auditors were compelled to get their meaning from the key words in any single sentence. When key words were defective the meaning of the sentence was lost to the auditors. When the speech error formed an entirely different but familiar word the auditors became confused because often it did not fit into the meaning to be conveyed by the sentence. Thus "black dog" was spoken by a pupil as *pak tok;* it was understood as "doctor" by five auditors who tried to fit the word into a context: "Bob had a little ———." The phrase "boys bought" was spoken as *poys*

pought and understood by some auditors as "Pass port," by others as "police pulled." The word "food" was spoken as *foot* and interpreted as "foot" or "suit" by several auditors. Often the word or phrase contained several types of errors, for instance: the phrase "for his dog" spoken *for ees tok* contains three errors, namely dropping of the aspirate *h*, substitution of the vowel *ee* for *–i–*, and dropping out the voice in both consonants in *dog* making it *tok*. The phrase was interpreted as "for Easter" by several auditors, which was a reasonable guess, but did not fit into the context and the entire meaning of the sentence was lost.

The consonants that gave the greatest trouble in this category were the stops *b*, *d*, and *g* and the affricative *j*. The continuitives *z* and *v* and *th* (thy) were less frequently malarticulated because in these sounds there is only a partial stoppage in the oral canal and the air is allowed to escape during the occlusion. Table 2 gives a list of the 10 sounds most frequently missed in this category.

In absolute frequency of errors (Table 1, Cols. 1) this category rank first and second at Mt. Airy and Clarke respectively. In terms of percent of possible errors (Table 1, Cols. 4) it ranks fourth at Clarke and second at Mt. Airy. The correlation between frequency of surd-sonant errors and speech intelligibility (Table 3) is —.49 and —.53 for the two schools.

2. *Substitution of One Consonant for Another*

The problem in consonant substitution appears to be one in which the deaf child substitutes a similar though perhaps easier sound for another. The question as to the relative difficulty of consonants will be discussed in a later section. The deaf child learns his consonants and vowels visually and tactually. Sounds which look alike are often substituted one for the other. In this study the sounds *r* and *s* rank highest in frequency of being replaced by other sounds. The sounds which were substituted for them most often were *w* for *r* and *sh* and *t* for *s*. There seems to be a fairly free interplay of substitution between *s* and *sh*. The Clarke pupils placed *s* and *sh* in second and third place respectively in frequency of substitutions, while Mt. Airy pupils placed them first and fourth respectively (Table 2).

Substitutions often lead to complete confusion of meaning. When *w* replaces *r* in words like "red," "ride," "row," words are formed which are totally dissimilar in meaning. The sentence "We rode

on a sleigh" was spoken as *we wote on a sleigh* and understood as "we went ———," "we wore ———." In the word "mother" *sh* replaced *th* and the word was heard as "muscle." The consonant *l* was replaced by *t* and "Polly" was heard as "putty." The fricatives *s* and *sh* may be replaced by sounds resembling a surd *y* as in the word "Sally" becoming *yalley*, or "ship" becoming *yip*, or "south" becoming *youth;* in such cases the consonant closures are not sufficiently constricted and the fricative sounds are not possible. In other cases the consonant occlusion becomes a complete closure, and *s* or *sh* are replaced by *t* or *ch*. Table 2 gives a list of the 10 sounds for which other sounds were most frequently substituted by both Clarke and Mt. Airy pupils.

Substitution as an error category ranks sixth in terms of frequency of errors in both schools. In terms of per cent of possible errors it ranks seventh. The correlation between frequency of substitution and speech intelligibility (Table 3) is —.28 (Clarke) and —.17 (Mt. Airy).

3. *Nasality*

Any consonant may become nasalized by the failure of the speaker to close the nasal pharynx by raising the velum. In the articulation of most of the consonants the quality and function of the sound depend upon a high buccal pressure, which is quite impossible if the air is allowed to escape through the nose. Only the consonants *m, n,* and *ng* in English are normally spoken with any degree of nasality. Deaf children often fail to close off the nasal pharynx in articulating consonants with the result that nasality predominates in both consonants and vowels. Lack of velum control is evidenced in both directions, namely, non-nasal consonants are often nasalized and nasal consonants often become complete stops. An example of both types of error is found in the phrase "at the corner" spoken *an the corter* and understood by several auditors as "on the porter." An example of failure to open the nasal pharynx for the proper articulation of the nasal sounds occurred in the word "summer" spoken as *sumber* and understood as "someboys." Teachers of the deaf, some of whom served as auditors in this study, become accustomed to excessive nasality in their pupils' speech and are able to understand it in spite of the defect. For this reason the frequency of this type of error is lower than one would expect if all excessively nasal sounds were tabulated.

TABLE 3

CORRELATION COEFFICIENTS: SPEECH INTELLIGIBILITY VERSUS FREQUENCY OF CONSONANT AND VOWEL ERRORS IN INDIVIDUAL CATEGORIES*

	Clarke			Mt. Airy		
Consonant Categories						
1. Surd-sonant vs. Intelligibility	—.49	*PE*	.05	—.53	*PE*	.05
2. Substitution vs. Intelligibility	—.28	*PE*	.07	—.17	*PE*	.06
3. Nasality vs. Intelligibility	—.34	*PE*	.06	—.19	*PE*	.06
4. Compound Consonants vs. Intelligibility	—.47	*PE*	.05	—.41	*PE*	.06
5. Abutting Consonants vs. Intelligibility	—.40	*PE*	.06	—.21	*PE*	.06
6. Arresting Consonants vs. Intelligibility	—.20	*PE*	.07	—.12	*PE*	.06
7. Releasing Consonants vs. Intelligibility	—.58	*PE*	.05	—.54	*PE*	.05
Total Consonant Errors vs. Intelligibility	—.70	*PE*	.04	—.71	*PE*	.03
Vowel Categories						
1. Substitution vs. Intelligibility	—.49	*PE*	.055	—.51	*PE*	.05
2. Diphthongs vs. Intelligibility	—.27	*PE*	.07	—.30	*PE*	.06
3. Diphthongization vs. Intelligibility	—.20	*PE*	.07	—.23	*PE*	.06
4. Neutralization vs. Intelligibility	—.36	*PE*	.06	—.39	*PE*	.06
5. Nasalization vs. Intelligibility	—.16	*PE*	.07	—.20	*PE*	.06
Total Vowel Errors vs. Intelligibility	—.61	*PE*	.05	—.56	*PE*	.05

*For the lay reader the "coefficient of correlation" may be defined as a statistical device used to express the degree and the nature of the relationship existing between two variables in quantitative terms. This relationship may be direct, or positive (when an increase in one variable, for instance, the age of children, is accompanied by an *increase* in another variable, the height of the children); or the relationship may be inverse, or negative (when an increase in one variable, for instance, an increase in the number of articulatory errors is accompanied by a *decrease* in the speech intelligibility score). The value of the coefficient may range up and down on a scale which extends from —1.00 through 0.00 to 1.00. A positive correlation indicates a *positive* relation, a zero correlation indicates an *absence* of relation, and a negative correlation indicates an *inverse* relation. The relationship is said to be "high" or "low" depending on the magnitude of the decimal fraction in the range between 0.00 and plus or minus 1.00. For correlation methods see (7, pp. 327 ff.)

Both continuants and stops are nasalized without very much discrimination by deaf pupils. Table 2 shows that among Clarke pupils *n*, *m*, *th* (then), *f*, and *p* were the highest group in terms of frequency of errors, while the Mt. Airy pupils had their greatest difficulty with *t*, followed by *th* (thy), *m*, *sh*, and *d*.

4. *Malarticulation of Compound Consonants*

A compound consonant is a group of two or more consonants fused to form a single consonant functioning within a single syllable either as a releasing or arresting consonant. The primary factor in the articulation of such consonants is a complete fusion of the component movements which make up the compound. This fusion is achieved in normal speech by making the multiple consonant movements as nearly simultaneous as possible. When the same articulatory member is employed in both components, as in *st*eal, complete simultaneity is impossible, of course, and the closure for *t* follows a momentary fricative phase for *s*. In the speech of deaf children two types of error occur relative to compound consonants: (*a*) the component movements are too slowly given and therefore fail to fuse; often the components occur as separate consonants with the indefinite vowel between them, thus increasing the number of syllables in the word. This type of error not only destroys the phonetic and acoustic form of the word in which it occurs but it also changes the rhythmic pattern of the phrase since adventitious syllables are added. An example of this type of error occurred in the sentence: "Did you brush your teeth this morning?" spoken as *did you burush,* etc. The sentence was understood as "Did you polish your teeth this morning?" The word "box" (boks) was spoken as *pokus,* and heard as "package" and "basket"; the auditors were completely confused by the two syllables in the word. Another example is that of the phrase: "birds fly south" spoken as *birdes f'y south* and heard as "thirty-five south." (*b*) A second type of error occurs when the pupils drop one or more components of the compound consonant. The word "fly" becoming *fy* in the illustration above is a case in point. In the sentence "Can you see the slate from there?" the *l* in slate was dropped. As a result the auditors wrote "shay," "show," "shave," and, less obvious, "chair."

Either of the components of a compound may be dropped; there seems to be no real trends in the examples studied. Compounds with *s* and *l* seemed to be more commonly malarticulated. Table 2 shows that both school groups missed *st* more often than any other combination. The data in Table 1 show that 30 per cent of all compounds attempted were mutilated by pupils in the two schools. The correlation between frequency of errors in this category and speech intelligibility is —.47 for Clarke pupils and —.41 for Mt. Airy (Table 3).

5. *Malarticulation of Abutting* (linking) *Consonants*

Abutting consonants are not to be confused with compounds. The latter belong to and function in a single syllable, while abutting consonants belong to and function in adjacent syllables. To illustrate: the compound *st* in the sentence "The star shines brightly" belongs to the syllable *star* and functions as a releasing consonant. On the other hand, the same pair of consonants which appear in the phrase "this tar—" belongs to adjacent syllables; the *–s* arrests the syllable *this* and the *t* releases the syllable *tar*.

The pupils tested in this survey had considerable difficulty with abutting consonants. The most frequent form of their difficulty appeared as an inability to make the transition from one consonant to the other. Instead of the normal linkage abutting pairs were separated by an indefinite vowel, thus "football" became *foo tu ball*, "on the first" became *o nu the first*. Both consonants of the abutting pair thus become releasing consonants and an adventitious syllable is added to the phrase. This type of error has the effect of distorting the acoustic form of the word and at the same time changes the rhythmic structure of the phrase. Teachers of the deaf evidently become accustomed to this type of error, for frequently they were able to interpret the meaning in spite of it. They were misled at times, however, especially if the combination seemed to form a familiar word or phrase. For instance, the phrase "it snows," spoken as *i tu snows* was interpreted as "Edison"; "popcorn" spoken as *po pu corn* was understood variously as "paper coins," "paper cows," "paper cloth"; "on Sunday," spoken as *o nu Sunday* was interpreted as "on a sailboat," "on a steam boat."

There were no special combinations forming abutting consonants that seemed to give more trouble than any others. In Table 2 the combination *–n:th–* leads the list for both school groups; there are insignificant differences, however, between this and other abutting pairs in the list. Correlations between frequency of errors in Category 5 and speech intelligibility was —.40 and —.21 for Clarke and Mt. Airy respectively (Table 3). In actual numbers this was the smallest category; in terms of per cent of possible errors it ranked sixth and fifth for the two schools respectively.

6. *Non-Function of the Arresting Consonant*

Any one of several things may happen to the arresting consonant in the speech of the deaf; it may be dropped completely, it may

become a releasing consonant in a following syllable, or it may lose all of its dynamic properties and become merely a passive oral gesture dangling at the end of the syllable. In rare cases a glottal stop was substituted for the final consonant. Words which end in fricatives such as *f* or *s* depend for their intelligibility upon the presence of the high pressure stream of air which accompanies the occlusion of these sounds. When the occlusion is incomplete or the orifice is opened too widely the arresting function is lost and the acoustic value of the consonant is either reduced or lost. To illustrate with actual examples taken from the records: In the phrase "ran a race" the *s* in race was non-functional, and the phrase became *ran a ra'*. Six auditors heard this as "ran away." The word "Otis" spoken as *Oti'* was heard as "oh dear." The phrase "big dog," spoken as *bi'do'* was understood by six auditors as either "book store" or "picture." "A pail," spoken as *a pai'* was heard as "all day," and "away."

From the point of view of the actual number of errors Category 6 ranked fourth for both school groups. In terms of per cent of possible errors it ranked second and third for Clarke and Mt. Airy respectively. It has the lowest ranking, however, when correlated with intelligibility (Table 3). This indicates that dropping the arresting consonant interferes less with speech intelligibility than any other defect studied in this survey. The consonants *l* and *t* lead the list in frequency of errors in this category (Table 2).

7. *Non-Function of Releasing Consonants*

The action of the releasing consonant in releasing the syllable consists of a momentary closure, either partial or complete, of the vocal canal, followed by a sudden opening. During the occlusion the air pressure builds up in the thoracic and buccal cavities and the sudden opening releases the syllable; thus the consonant movement is heard largely by its effect upon the vowel. A complete co-ordination between breathing muscles which move the air column and the consonant movement which closes the oral cavity is of vital importance in this process. This coördination is often lacking in the speech of deaf children. The articulatory movements may be too slow or fail to develop the necessary pressures, indicating that the articulatory movements and the movements of the breathing muscles are inaccurately timed. Again the movements may be incomplete leaving the orifice too large in case of the continuants and preventing

the development of the proper degree of pressure. In either case the effect upon the auditor is that of having the syllable begin with a vowel.

Table 2 shows the 10 consonants most frequently dropped as releasing consonants (Category 7). The aspirate *h* heads the list for both schools; *l*, *r*, and *y* follow in the same order. The aspirate *h* perhaps should be put into a separate category since it is not, strictly speaking, a consonant. It was included, however, for simplification. The aspirate involves a sudden aspiration of air through the partially open glottis immediately preceding the vowel. The mechanism for its production involves a separation followed by an approximation of the vocal cords during a single syllable. No other organ of articulation is involved and the conformation of the oral cavity takes on the form necessary for the production of the following vowel. Instead of obstructing, or partially obstructing the vocal canal, therefore, the aspirate opens the glottis allowing for a freer flow of air. The deaf pupils tested in this survey had difficulty in making the glottal adjustment necessary for the aspirate.

The other consonants most frequently dropped in this category, *l*, *r*, *y*, *th*, *s*, are for the most part continuants which involve articulatory movements of delicate adjustment to and from an opposing surface, but which result in an incomplete closure of the oral canal leaving an orifice through which air under pressure escapes producing high frequency fricative sounds. The accuracy of the movement, and the degree of adjustment of the orifices determine the identity of the consonants. The complete stops are less frequently dropped from the releasing position.

A few illustrations will indicate the variety of defects and the resulting confusion of the auditors. Example of dropping the aspirate *h* occurred in the phrase "did you hear the ———" spoken as *did you 'ear the* ———, and understood as "did you ever ———"; "Mary has," spoken as *Mary 'as* and understood as "Mary is"; "for his dog" spoken as *for ees tok*, understood as "for Easter." The phrase "go for a ride" was spoken with several errors, among them dropping the *g*, and spoken as *o-e for a wite;* this was understood by several auditors as "Oh, for a while"; "your name" was spoken as *'our name* and understood as "our name." When the closure for *sh* was incomplete the resulting sound was often mistaken for the aspirate; "She went for a ride" with the *sh* dropped was heard as *he went for a ride*.

Category 7 shows the highest correlation with speech intelligibility (Table 3). It ranks highest in actual number of errors for Clarke pupils, and second largest with those of Mt. Airy. It ranks third and fourth highest in terms of per cent of possible errors for Clarke and Mt. Airy respectively.

C. Analysis of Individual Vowel Error Categories

The experimenters were less critical in determining the degree of vowel accuracy than that of consonants. The justification for this lies in the nature of the vowel itself and in degree of tolerance of the normal ear for a wide degree of vowel distortion. Classifications of vowels into rigid categories in which each vowel is assigned a definite oral conformation and definite bands of frequencies representing a distinctive quality can be done only by having subjects intone the vowels singly and without context. Vowels in normal speech become mere approximations of these rigid forms. Speech is intelligible and may be even considered normal in which the vowels only remotely approach the degree of accuracy set up by the phoneticians and phonologists. The acoustic structure or quality of vowels varies widely within a group of individuals. Furthermore, the immediate phonetic context, the degree of stress and rate of syllable utterance all are factors modifying the vowel.

Vowels in the speech of the deaf rarely reach the degree of accuracy attained by those in the speech of normal hearing persons. Vowel errors were listed only when the assigned vowel was totally unrecognizable. The various types of vowel errors which occurred in the speech of the pupils studied in this survey are discussed below. Approximately 11 per cent of all vowels spoken by pupils of the two schools were listed as errors.

1. *Vowel Substitution*

Deaf children not only substitute vowels whose formations are similar, for instance, *–i–* (bit) for *ee* (beet) but they also substitute one vowel for another in which the formations are dissimilar. The reason for this lies perhaps in the fact that they have only the rather vague movement patterns as cues. Illustrations from the records will indicate the effects of vowel substitution upon speech intelligibility. In the sentence "The cow was in the barn," the vowel in "barn" was changed so the word became *ban*, this was understood as "bank" by several auditors; the vowel in "made" was changed

to *ee,* one auditor heard "need," another "meat." The vowel as well as the consonant in "buy" was changed and the word became pay, this was heard as "pate," "paints," "please," "bring," and "he." The phrase "my team" was spoken as *ma tam* and was understood by four auditors as "madam." Many vowel substitutions occurred which made the word totally unfamiliar to the auditors and no interpretations were given.

Vowel substitution was the largest category of errors for both schools. According to Table 1 it makes up 55 per cent of all vowel errors for Clarke and 66 per cent of all vowel errors for Mt. Airy. Other vowels were substituted for more than 6 per cent of all vowels attempted by pupils of both schools. Table 2, Category 1, shows that Clarke pupils had trouble with diphthongs *a-e* (day) and *i-e* (pie) for these head the list of five most frequently substituted vowels. Mt. Airy on the other hand, had more trouble with the short vowels *–ĭ–* (bit) and *–e–* (bet). The correlation between frequency of vowel substitution and speech intelligibility was —.49 and —.51 for Clarke and Mt. Airy respectively (Table 3).

2. *Errors Involving Diphthongs*

A diphthong, properly articulated, is a fusion of two movements; it is similar in this respect to compound consonants. It is interesting to note in this connection that the same types of errors occurred, namely, the two components were separated or prolonged until two distinct vowels were heard, as *baw-ee* for *boy,* or one component usually the final one, was dropped, as *mah* for my.

There are numerous examples of auditor errors directly traceable to the error of making two syllables out of diphthongs. For instance, in the phrase "I have ———," spoken as *ah-ee have* the auditor heard "Annee has," two others wrote nothing. Again, the phrase "our cook" spoken as *o-wer cook* was understood as "how do" "hello," and "Helen's cook." In each of these cases the auditors heard the two syllables and tried to approximate the sounds with familiar words. Examples of dropping the final member of the diphthong occurred in the sentence "Have you saved some money?" spoken as ——— *saft some money,* the verb was unrecognized by two auditors, heard as "seen" by two others, and as "each" by one. The diphthong *i-e* (die) was frequently repeated as *ah,* and understood at times by auditors but usually a word more nearly like the one actually spoken was substituted.

The diphthong *i-e* was most frequently mispronounced by the pupils of both schools. Only nine per cent of all vowel errors were of this type; likewise nine per cent of all diphthongs were malarticulated by pupils of the two schools. Correlation between diphthong errors and intelligibility (Table 3) is —.30, and —.27 for Mt. Airy and Clarke respectively.

3. *Diphthongization of Pure Vowels*

One of the chief characteristics of the speech of the profoundly deaf children is a lack of coördination between the articulatory organs and the breath pulses. Instead of the accurate timing of consonant or vowel movements and syllable pulses which normally occurs in the speech of normal speakers, deaf children appear to make slow transitory movements during the articulation of the vowel. The continuation of the stream of voice during these transitory movements renders the vowel as a diphthong. At times the transition between vowel and arresting consonant or between releasing consonant and the vowel are much too slow, thus the vocal stream undergoes modifications which are totally unlike those occurring in the normal stream of speech. To illustrate, instead of "how do you do" the deaf child may say *how-ee do-ee you-ee do-ee.* In this particular example the transition from the vowel *oo* to the following consonants *d, y,* and *d* are so slowly executed that the transitory movements are heard and appear as a prominent part of the vowels. The failure to arrest the stream of voice during the transition or to speed up the transitory movements makes the movement audible. In the sentence "Who gave you the ring?" the word *who* was spoken as *whoo-ee* and understood by one auditor as "Lewis" which was a good guess. "Where are my shoes?" became *Where are my shoo-is!* The last word was heard as "Julia." "Bread" was spoken as *bre-ud,* and understood as "we had," "will you," "do you."

This was one of the less important of the vowel categories. Diphthongization of the vowels o^2o (book) and o^1o (boots) occurred most frequently in this category. Correlations between number of errors and intelligibility were small, probably negligible. Nine per cent of all vowel errors were of this type, and 1.3 per cent of the total errors possible were made by Clarke pupils. Mt. Airy pupils made fewer errors in this category.

4. *Neutralization of Vowels*

Neutralization of vowels is a form of vowel substitution; it is listed as a separate category because it represents a very definite type of error rather than promiscuous substitution.

A vowel is neutralized when a minimum modification occurs in the vocal canal during its production. The quality and often the duration of such vowels are similar to those of short unaccented vowels. There is a tendency, even in cultivated speech, for vowels in unaccented syllables to become neutral vowels. In slovenly speech this process may reach a point where entire syllables are dropped. Such speech, however, may still be intelligible to those familiar with it. When syllables usually accented lose their accent and vowels become neutralized the speech becomes more or less unintelligible. Vowels in the speech of deaf children, both in accented and unaccented syllables, are often neutralized. In the sentence "Will you be home tonight?" the subject neutralized the vowel in "will" and said *wul*. Two auditors in six understood the sentence; the remaining four wrote "—— be home tonight." The same pupil in repeating the word "train" said *trun*. Not one of the six auditors understood this word; there were, however, other errors in the sentence. In the phrase "She gave me ——" the vowel in "me" was spoken as *muh;* two auditors understood the word possibly with the aid of its context; six others either left blanks for the word or wrote "one" instead of "me." In the sentence "A baby bird ——" the pupil said *bud*, but since this did not fit in with the context most of the seven auditors wrote "A baby ——," two wrote the word "good" for "bird."

Neutralization represents the second largest category of vowel errors. Nineteen per cent of all vowel errors were of this type for Clarke pupils and nine per cent for Mt. Airy. The correlations between frequency of neutralization of vowels and speech intelligibility (Table 3) are significant: Clarke —.36 and Mt. Airy —.39.

5. *Nasalization of Vowels*

. One of the characteristics of English vowels is the predominance of oral resonance; there are no nasal vowels in the language. Excessive nasality, therefore, may be considered as a defect.[8] The speech

[8]The term nasality as used in this paper means the effects produced by failure of the velum to close the nasal pharynx.

of many of the deaf pupils studied was effected by nasality in varying degrees. When this reached what the experimenters considered an excessive degree the sounds were listed in the nasality category. Vowels occuring in syllables with nasalized consonants were more commonly nasalized. Those pupils, therefore, who were inclined to nasalize consonants also were the chief offenders in nasalizing vowels. It is difficult to determine whether the audition-errors caused by nasality are to be attributed to the defective vowel or consonant since both result from the lack of control or malfunctioning of the velum.

Table 2 shows that there is little agreement between the pupils of the two schools as to vowels most frequently nasalized. Indeed, the consonants associated with the vowel probably determine whether or not it will be nasalized rather than the characteristics of the vowel itself.

The correlations between nasality of vowels and speech intelligibility were too low to be significant (Table 3). This does not mean that this type of defect is unimportant or that it may be ignored in speech correction. It is so intimately related with nasality of consonants, however, that it is difficult to evaluate it separately. Correction of nasality of consonants by developing the proper control of the velum will affect the vowels as well. The problem of nasality, therefore, unlike the other types of errors studied in this paper, is common to both consonants and vowels.

Consonant errors are more frequent and are apparently more important for speech intelligibility than vowel errors. Table 3 shows coefficients of correlation between speech intelligibility: and (*a*) individual error categories; (*b*) total consonant errors; and (*c*) total vowel errors. The data from the two schools are in close agreement. The differences between consonants and vowels are significant. These data will be analyzed further by methods of partial correlation and discussed in Section G, pp. 354-356.

D. Relative Importance of Individual Error Categories for Speech Intelligibility

Pupils of two schools for the deaf, using similar test material, have been found to make similar articulatory errors. Clarke pupils mispronounced 22 per cent of all consonants and 12 per cent of all vowels attempted, while Mt. Airy pupils mispronounced 20 per cent of all consonants and 11 per cent of vowels. Speech intelligi-

bility scores were determined for each pupil based upon these speech tests. Total consonant errors have been classified into seven error categories and vowel errors into five categories. A practical question arises as a result of the classification, namely: Is the intelligibility of the speech of these pupils affected to a greater degree by one type of error than by another? A positive answer to this question would indicate that teachers would be justified in spending relatively more time and effort correcting and training children to avoid those errors which interfere to a greater degree with speech intelligibility.

There are two criteria available for determining the relative importance of the several error categories: (*a*) The degree of correlation between the number of errors made by individual pupils in each category and the speech intelligibility score; and (*b*) the ranking of the individual categories on the basis of *per cent of possible errors.*

1. *Correlations*

A correlation coefficient of significant magnitude between the number of errors made by each pupil and speech intelligibility means that intelligibility is related to and significantly varies with frequency of errors. A negative correlation between these two variables means that as the frequency of errors increase the intelligibility scores decrease. It is possible, therefore, to use the coefficient and its probable error as a criterion of the relative importance of each type of error. A list of coefficients of correlation with their probable errors is presented in Table 3. There is a striking similarity between the two schools in degree of correlation. The rank order of the several error categories, as determined by the coefficients of correlation is identical. The three highest ranking consonant categories are: 7 (dropping releasing consonants), 1(surd-sonant errors), and 4 (errors of compound consonants). The coefficients of correlation for these three categories fall between —.47 and —.58 for Clarke pupils, and between —.41 and —.53 for Mt. Airy. The differences between them are not statistically significant. The other four categories are ranked as follows: Category 5 takes the fourth rank, 3 the fifth, 2 the sixth, and 6 the seventh. The coefficients are smaller for these four categories; those for Clarke pupils are significant with the exception of Category 6. The correlations for Mt. Airy pupils are rather low for the four categories and are probably insignificant.

The following conclusions may be drawn on the basis of this criterion: (*a*) Consonant errors which have been classified under Categories 7, 1, and 4 have an equal effect upon the intelligibility of the speech of the pupils studied; (*b*) the magnitude of the coefficients of correlation indicate that these three types of errors are of considerable significance in rendering speech unintelligible; (*c*) the remaining four categories are of lesser importance but they also contribute to the unintelligibility of the speech of deaf children.

Table 3 also shows the coefficients of correlation between the *vowel error categories and speech intelligibility.* Again the rank order is identical for the two school groups. Category 1 (substitution) appears as the most important with Categories 4, 2, 3, and 5 following in that order.

2. *Per Cent of Possible Errors*

As a second criterion of the relative importance of the several categories the frequency of each category expressed in terms of *per cent of possible errors* may be used. There is some justification for the use of this percentage rather than the absolute number of errors as a criterion of relative importance. The number of errors that any subject may make is limited by the number of consonants or vowels in the speech material which may fall into each of the several categories. Since this number of potential errors varies with each error category a percentage representing the ratio of actual and potential errors becomes the most adequate expression of this relationship. For instance, Table 1, Column 1, shows that (Clarke) Categories 1 and 4 are almost equal in absolute number of errors. The number of errors possible in these two categories, however, as shown in Column 3 (Table 1), is 6,352 and 920 respectively. It appears, therefore, that the 399 errors in Category 1, and the 355 in Category 4 are not directly comparable, since the chances for error are 7 to 1 in favor of Category 1.

On this basis of comparison Category 4 (compound consonants) is by far the outstanding one in the data from both school groups (Figure 2). The four highest ranking categories are 4, 6, 7 and 1, Clarke, and, with a slight shift in rank order, 4, 1, 6 and 7, Mt. Airy. The remaining categories follow in an identical order for the two schools: Category 5 takes the fifth rank, 3 the sixth, and 2 the seventh. There is a striking agreement in the two sets of criteria both between the two school groups using either criterion

alone, and between the two criteria themselves with the single exception of Category 6 which ranks seventh in Criterion 1 and second in Criterion 2. No apparent explanation is available for this discrepancy.

With reference to vowel errors the second criterion of relative importance places Category 2 (diphthong errors) in first rank, with Category 1 (substitution) second. The percentages for the remaining three categories are relatively small.

Both criteria of relative importance of the several types of consonant errors identify compound consonant errors, surd-sonant errors, and errors involving releasing consonants as being of high importance relative to speech intelligibility. The agreement of independent criteria in ranking the remaining four types, or categories of consonant errors is close with the exception of Category 6 (dropping the arresting consonant). The rankings of the five types of vowel errors show less agreement than that of the consonants.

E. Analysis of Articulatory Errors According to Degree of Hearing Loss

The data presented in this study confirm the already well-known observation that the presence of even a small amount of residual hearing may exert an important effect upon the speech development of deaf children. The more recently extended use of electrical hearing aids in the education of deaf children has further increased the benefits derived from residual hearing. Some of the pupils in both schools included in the present study have had the advantage of hearing aids in their speech development and in general education. Mt. Airy, however, at the time of this study, had a larger number of pupils who were spending the entire school day in classrooms equipped with hearing aids. The advantages derived from these hearing aids are reflected in the data to be presented in this section.

Audiograms were available for all the pupils studied. It was possible, therefore, to analyze the number of speech errors in terms of the degree of hearing loss. The method of classifying deaf children according to degree of hearing loss presented by Guilder and Hopkins (6) was used. This method while arbitrary is based upon quantitative scores derived from audiometric tests covering a range of eight octaves. Three classifications are used: Group *A* is made up of hard-of-hearing children who acquire speech largely through hearing but are not able to keep up with public school

children because of hearing defects. Group *B* is made up of partially deaf children who are not able to acquire speech in the normal manner but whose speech can be benefited by electrical hearing aids. Group *C* is composed of profoundly deaf children who have little or no usable residual hearing. A more rigid classification would take into account the contour of the audiometric curve over the eight octaves as well as the decibel level. This would take into account high frequency, low frequency, or median frequency deafness. No standardized classification method of this type is available, however, at present.

The 87 Clarke pupils were divided as follows: Group *A*, 11 pupils; Group *B*, 20 pupils; and Group *C*, 56 pupils. A similar classification of 105 Mt. Airy pupils shows Group *A*, 10 pupils; Group *B*, 36 pupils; and Group *C*, 59 pupils. Groups *A* and *C* are similar in number for both school populations, while Group *B* in Mt. Airy is nearly double that of Clarke.

The data derived from these groups were analyzed from three different aspects and presented in Tables 4 and 5; these show: (*a*) The average number of consonant and vowel errors per pupil; (*b*) frequency of errors in terms of *per cent of total errors made;* and (*c*) frequency of errors in terms of *per cent of possible errors.*

Figure 3 is a graphic representation of the totals given in Tables 4 and 5; it represents the average number of consonant and vowel errors per pupil for the three degrees of deafness. These averages increase significantly with increase of hearing loss. Columns representing Groups *C* are almost identical in height for the two school groups; those representing Groups *A*, and to an even greater extent, Groups *B* show a considerable larger number of consonant errors for the Clarke pupils. The larger number of Mt. Airy pupils in Group *B* might conceivably account for the difference. A more plausible explanation is suggested, however, by the fact that Mt. Airy pupils in Groups *A* and *B*, at the time of this study, had had the advantages of hearing aids in all classroom work for periods of two to five years.

The differences in number of articulatory errors are reflected in the intelligibility scores of Groups *A* and *B*, Mt. Airy. This is partly obscured by the fact that the general intelligibility level is higher for Clarke pupils than that for Mt. Airy pupils (Table 4). The intelligibility scores of Mt. Airy pupils in Groups *A* and *B*, however, are almost identical with those of the corresponding Clarke

TABLE 4

THE ANALYSIS OF CONSONANT ERRORS ACCORDING TO DEGREE OF HEARING LOSS

	Clarke								
Degree of deafness	*A.* (11 pupils)			*B.* (20 pupils)			*C.* (59 pupils)		
Error categories	Av. per pupil	% of total	% of poss. errors	Av. per pupil	% of total	% of poss. errors	Av. per pupil	% of total	% of poss. errors
1. Surd-sonant	1.2	14.3	1.6	4.0	17.5	5.3	5.5	21.8	7.5
2. Substitution	.6	7.7	.6	1.0	5.1	1.1	2.0	7.4	1.8
3. Nasality	.3	2.9	.3	3.0	12.5	2.8	3.0	11.7	2.9
4. Compound consonants	2.0	25.0	20.0	4.0	18.4	39.0	4.4	17.4	42.0
5. Abutting consonants	.4	4.4	1.6	.5	2.5	2.4	2.0	7.4	8.1
6. Arresting consonants	3.0	33.0	7.6	4.0	18.8	11.6	2.7	10.6	7.4
7. Releasing consonants	1.0	12.0	1.5	5.0	22.8	7.8	6.0	23.6	9.1
Total	8.5	100.0	8.1	21.5	100.0	22.0	25.4	100.0	25.0
Average intelligibility	78 (range 70-96)			44 (range 20-76)			21 (range 0-50)		
	Mt. Airy								
	A. (10 pupils)			*B.* (36 pupils)			*C.* (59 pupils)		
1. Surd-sonant	1.5	22.0	2.0	6.0	37.0	7.8	10.5	41.0	14.3
2. Substitution	1.0	15.0	1.0	1.0	10.0	1.6	1.8	7.0	1.8
3. Nasality	.1	1.5	0.0	1.6	10.0	1.6	2.1	8.3	2.1
4. Compound consonants	1.8	26.0	17.0	2.3	14.0	22.0	3.2	12.5	30.5
5. Abutting consonant	.6	8.7	2.6	.7	4.3	3.0	1.0	3.6	4.0
6. Arresting consonants	1.6	23.0	4.5	2.4	15.0	8.5	2.4	9.4	6.7
7. Releasing consonants	.3	4.3	.5	1.6	9.7	2.6	4.6	18.0	7.1
Total	6.9	100.0	6.8	16.0	100.0	15.7	25.3	100.0	25.0
Average intelligibility	69 (range 34-93)			43 (range 3-87)			9 (range 0-30)		

TABLE 5
THE ANALYSIS OF VOWEL ERRORS ACCORDING TO DEGREE OF DEAFNESS

					Clarke				
Degree of deafness	*A*. (11 pupils)			*B*. (20 pupils)			*C*. (56 pupils)		
Error categories	Av. per pupil	% of total	% of poss. errors	Av. per pupil	% of total	% of poss. errors	Av. per pupil	% of total	% of poss. errors
1. Substitution	.6	50.0	.6	3.8	53.0	2.1	6.3	56.0	8.0
2. Diphthongs	.2	20.0	2.0	.5	6.1	4.9	1.1	9.7	11.8
3. Diphthongization	.2	20.0	.4	.5	6.1	.6	1.1	10.2	1.6
4. Neutralization	.1	10.0	.1	2.0	27.0	2.5	2.0	17.7	2.5
5. Nasality	0.0	0.0	0.0	.5	7.5	.7	.7	6.0	.9
Total	.9	100.0	1.2	7.3	100.0	9.2	11.2	100.0	14.2
					Mt. Airy				
	A. (10 pupils)			*B*. (36 pupils)			*C*. (59 pupils)		
1. Substitution	2.5	80.0	3.1	3.1	72.0	3.2	6.4	63.0	8.1
2. Diphthongs	.3	10.0	3.0	.4	8.3	2.4	1.2	12.0	13.7
3. Diphthongization	.2	6.0	.3	.1	2.6	.2	.5	5.4	.8
4. Neutralization	.1	3.0	.1	.3	7.0	.3	1.0	9.2	1.2
5. Nasality	0.0	0.0	0.0	.4	10.3	.5	1.0	9.6	1.2
Total	3.1	100.0	4.0	4.3	100.0	4.4	10.0	100.0	12.8

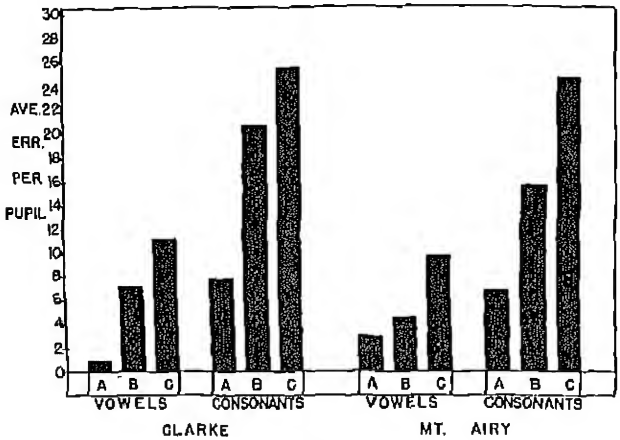

FIGURE 3

AVERAGE NUMBER OF CONSONANTS AND VOWEL ERRORS PER PUPIL ARRANGED ACCORDING TO THREE DEGREES OF HEARING LOSS

A—Hard-of-hearing pupils who acquire speech by means of their defective hearing mechanism.

B—Partially deaf pupils who have a usable amount of residual hearing, but do not acquire speech by ear except by the use of hearing aids.

C—Profoundly deaf pupils with little or no residual hearing.

groups. The great difference in intelligibility, therefore, between Clarke and Mt. Airy pupils lies in the speech of the profoundly deaf, Groups *C*, where the Clarke and Mt. Airy averages were 21 and 9 respectively. It might be suggested that the higher level of average intelligibility at Clarke is due to the fact that the Clarke pupils are required to use speech to a greater degree outside of classroom situations. Whatever the cause of the higher level of intelligibility at Clarke, the use of hearing aids with something like 44 per cent of Mt. Airy pupils (Groups *A* and *B*) tends to compensate for other advantages held by the Clarke pupils.

When the total number of articulatory errors for pupils in the three divisions of hearing loss is broken up into error categories and expressed in terms of *per cent of total errors* (Columns 2, Tables 4 and 5) some interesting differences between the several categories appear. Figure 4 shows these differences graphically. Categories 1, 3, and 7 (Consonants) show a distinct increase in frequency with increased hearing loss. This means that as the degree of deafness

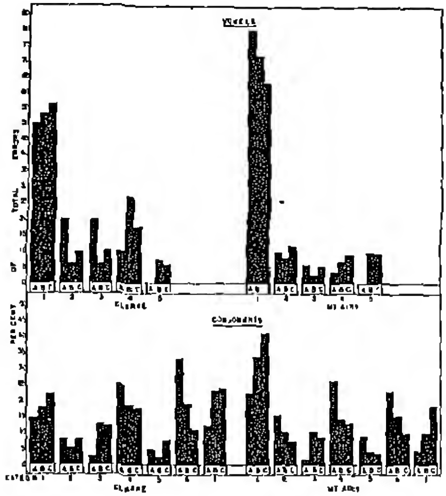

FIGURE 4

CONSONANT AND VOWEL ERRORS IN TERMS OF PER CENT OF TOTAL ERRORS, ARRANGED ACCORDING TO: (*a*) ERROR CATEGORIES; AND (*b*) DEGREE OF HEARING LOSS

Consonant Categories 2, 4, and 6 decrease in height with increase in hearing loss; this is a reversal of the trend in Categories 1, 3, and 7 (cf. Figure 3).

Categories:

1. Surd-sonant
2. Substitution
3. Nasality
4. Compound consonants
5. Abutting consonants
6. Arresting consonants
7. Releasing consonants

1. Substitution
2. Diphthongs
3. Diphthongization
4. Neutralization
5. Nasality

increases a greater portion of total errors fall into these categories. Categories 2, 4, and 6, on the other hand, show a complete reversal of this trend. This means that although the frequency of total errors increases with an increase in hearing loss, a greater number of consonant errors made by hard-of-hearing and partially deaf pupils fall into Categories 2, 4, and 6. Category 5 shows no clear trend. These facts are further confirmed by coefficients of correlation between number of errors in individual categories and auditory scores for individual pupils. Coefficients of correlations for six consonant error categories taken from data covering 192 pupils appear in

TABLE *B*

	Category				
Surd-sonant	1	(Errors per pupil vs. auditory score)	—.43	*PE*	.04
Substitution	2	(Errors per pupil vs. auditory score)	—.18	*PE*	.05
Nasality	3	(Errors per pupil vs. auditory score)	—.17	*PE*	.05
Compound consonant	4	(Errors per pupil vs. auditory score)	—.18	*PE*	.05
Arresting consonant	6	(Errors per pupil vs. auditory score)	.00	*PE*	—
Releasing consonant	7	(Errors per pupil vs. auditory score)	—.48	*PE*	.04

Table *B*. The relatively high correlation with auditory scores in Categories 1 and 7 indicate the close relationship between these types of errors and degree of deafness. (The correlation for Category 3, which also follows this trend in Figure 4, is unaccountably lower.) Categories 2, 4, and 6 which show a decrease in per cent of total errors with increased hearing loss (Figure 4) have very low correlations with hearing scores, indicating that these error types are dependent to a lesser degree upon hearing.

These facts indicate that the sonant-surd distinction, releasing consonants (especially *h, l, y,* and *r,* Table 2) and to a lesser degree the control of the velum (nasality) are affected by residual hearing to a relatively greater degree than the other categories. On the other hand, it means that consonant substitution, compound consonants, and arresting consonants are effected to a lesser degree by residual hearing.

The above statements do not mean that partially deaf children actually make more articulatory errors in some of the categories than those who are profoundly deaf. There is always a greater absolute *number* of errors to be found in the speech of pupils in Group *C* (Figures 3 and 5). The differences evident in Figure 4, however, show that error Categories 2, 4, and 6 make up a relatively larger proportion of total errors in the speech of Groups *A* and *B*. The fact that these relationships are more clearly defined in the Mt. Airy data (Figure 4) where hearing aids are more widely used is further evidence for the reality of the differences.

A possible explanation of these differences is offered by: (*a*) inherent differences between the types of consonants appearing in the error categories in question, namely, acoustic and motor-phonetic differences; and (*b*) differences in the manner in which speech is

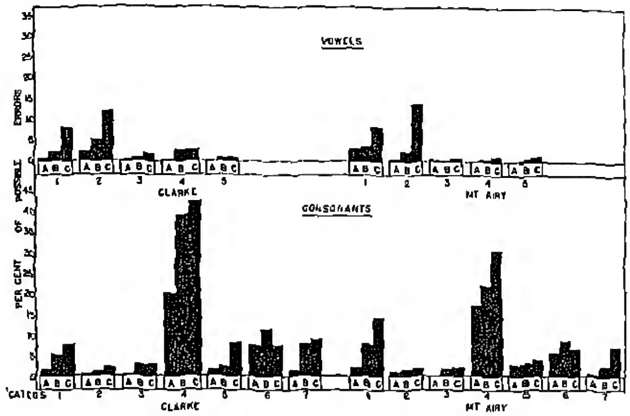

FIGURE 5

CONSONANT AND VOWEL ERRORS IN TERMS OF PER CENT OF POSSIBLE ERRORS ARRANGED ACCORDING TO: (*a*) ERROR CATEGORIES; AND (*b*) DEGREE OF HEARING LOSS

The more profoundly deaf pupils make a greater number of errors in all categories.

Categories:

1. *Surd-sonant*
2. Substitution
3. Nasality
4. Compound consonants
5. Abutting consonants
6. *Arresting consonants*
7. Releasing consonants

1. *Substitution*
2. Diphthongs
3. Diphthongization
4. Neutralization
5. Nasality

learned by children with different degrees of hearing loss. Pupils in Group *A*, and to a lesser degree, those in Group *B* acquire speech partly by means of their defective hearing mechanism. Articulatory errors are to be expected more frequently, therefore, in those speech sounds which are acoustically similar, or in those which provide only slight auditory cues. The errors involved in substituting *s* for *sh* or *th*, *th* for *f*, *w* for *r*, and even *t* for *p* or *k* clearly fit this description. Likewise, the *s* in the compound *st*–, the *l* in *pl*–, and the final consonants in words such as *bit*, *cook*, *fish*, and *this* provide very slight auditory cues especially for pupils with median hearing losses.

Pupils in Group *C*, on the other hand, never hear any speech sounds. They learn to build words and phrases by "stringing together" a series of individual articulatory movements. The relative acoustic value of different sounds is not, therefore, a problem. Their

problem is rather one involving the relative difficulty or complexity of individual articulatory movements, and of coördinating concurrent articulatory and breathing movements in repeating a series of syllables. The articulatory movements in producing sonant stops (Category 1) are undoubtedly more difficult than those involved in the surd stops (11) (Table 15). Likewise the action of the velum (Category 3) which must be coördinated with tongue, lips, and jaw movements in all consonant articulation adds to the complexity of articulatory movements. Again, the delicacy of adjustment required of articulatory organs and breathing mechanism in the production of syllables released by the aspirate *h* and by the consonants *l, y,* and *r* (Table 2, Category 7) is relatively more difficult than that of sounds involving greater buccal pressures and greater degree of occlusion (see Table 15 for order of difficulty of consonants). Hence, profoundly deaf children who learn speech solely as a series of consciously acquired motor responses make a relatively larger percentage of errors in those consonants which involve a greater complexity of motor coördination in their production.

The percentages of total vowel errors for the three degrees of hearing loss (Figure 4) show no significant trends. Vowel substitution makes up the larger proportion in all three groups. Mt. Airy pupils show a decrease in vowel substitution with increase in hearing loss. The actual number of vowel errors, however, made by Groups *A* in both schools are too small for a reliable analysis. Categories 2 and 3 in the Clarke data show apparent decreases with increase of hearing loss, but here again the number of errors for Group *A* is too small.

Figure 5 shows the articulatory errors plotted in terms of per cent of possible errors at each category for the three degrees of deafness. When plotted from this aspect the data in all categories show an increase in per cent of possible errors with increase in degree of hearing loss. In other words, as hearing loss increases a greater number of errors occur. Categories 1, 4, and 7 are outstanding in this respect. Categories 2, 3, 5, and 6 show minor differences between Groups *A, B,* and *C.*

Summary

Articulatory errors and audiometric scores are definitely correlated; the relationship is an inverse one. A more detailed analysis

in which errors are divided into separate categories, however, qualifies this relationship. There are clear indications that some types of articulatory errors are more closely associated than others with the degree of hearing-loss. Consonant substitution and errors involving compound and arresting consonants appear in greater proportion among the hard-of-hearing and partially deaf pupils than among the profoundly deaf. Failure to make the surd-sonant distinction, excessive nasality, and dropping the releasing consonant, especially the aspirate *h* and consonants *l*, *r* and *y*, are more highly correlated with degree of hearing-loss. Suggested explanation for these differences takes into account: (*a*) the acoustic and motor-phonetic differences in the consonants involved in the different error categories; and (*b*) differences in the manner in which speech is learned by pupils with different degrees of hearing loss.

F. Analysis of Articulatory Errors According to Age

An analysis of the speech data on the basis of age should show two things: (*a*) Whether or not there is an increase or a decrease in articulatory errors and speech intelligibility with increasing age (years in school); and (*b*) whether or not certain types of speech errors occur more frequently than others at any given age level.

The data were analyzed separately for the two school groups. The ages of the pupils studied ranged from 8 to 20 years. They were divided into two-year age groups according to their nearest birth dates. The following age-groups were used: 8-9, 10-11, 12-13, 14-15, 16-17, and 18- years. The number of pupils in each age-group for the two schools taken separately ranged from 15 to 24 with the following exceptions: Clarke age-group 8-9, 1 pupil; 10-11, 8 pupils; Mt. Airy age-group 8-9, 6 pupils. The reliability of the averages is doubtful in the 8-9-year groups for both schools and in the Clarke 10-11-year group.

A summary of the articulatory errors according to age is presented graphically in Figure 6. The columns represent the average number of consonant and vowel errors per pupil in all categories for each age-group. The average number of consonant errors appears to increase with age from the 10-11-year age-groups upwards in both school groups. The peak is reached at the 14-15-year level for Mt. Airy pupils, but for the Clarke pupils the increase continues throughout the age range. Vowel errors show very little change with age.

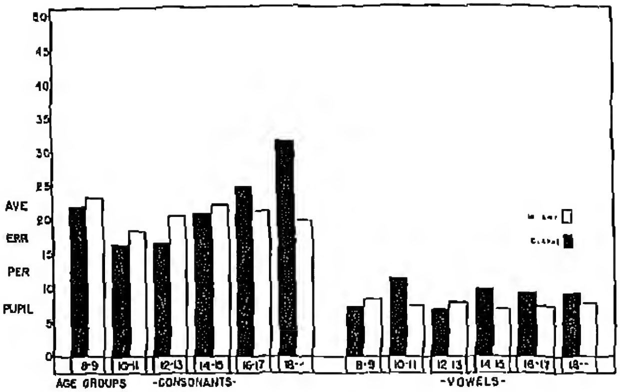

FIGURE 6

THE AVERAGE NUMBER OF CONSONANT AND VOWEL ERRORS PER PUPIL ARRANGED ACCORDING TO TWO-YEAR AGE GROUPS

Consonant errors tend to increase with age especially among Clarke pupils; vowels remain relatively constant. (Clarke age-group 8-9 represents only one pupil.)

The apparent increase in consonant errors with age means that the speech of the older pupils is less intelligible since intelligibility is negatively correlated with frequency of consonant errors (Table 3). The average intelligibility scores for pupils in the several age-groups (Table 6) show this fact. The Clarke intelligibility averages increase up through the 12-13-year level, then rapidly decline. The Mt. Airy averages show minor variations but no definite trend with age.

An analysis of the individual error categories according to age is presented in Table 6. Differences appear not only in the trends of the individual error categories with age, but also between the averages of the two school groups. These differences are more clearly defined in the graphic presentation of the consonant categories in Figure 7. The columns represent average number of errors per pupil in each error category. The graph represents trends rather than statistically significant differences. The following trends appear:

Surd-sonant errors (Category 1, Figure 7) show an increase in individual averages with age in both school groups. The greater number of errors of this type made by Mt. Airy pupils is indicated

TABLE 6

AVERAGE CONSONANT AND VOWEL ERRORS PER PUPIL IN EACH CATEGORY ACCORDING TO AGE; AVERAGE INTELLIGIBILITY SCORES AT TWO-YEAR AGE LEVELS

	Clarke						*Mt. Airy*					
Age groups	8-9	10-11	12-13	14-15	16-17	18-	8-9	10-11	12-13	14-15	16-17	18-
Consonant Categories												
1	10.0	3.6	3.1	4.3	5.3	5.1	5.3	5.7	7.6	9.8	9.0	8.6
2	2.0	1.4	1.6	2.4	1.7	.9	1.7	1.5	1.5	1.7	1.7	1.8
3	3.0	1.9	1.0	2.0	2.8	5.1	5.3	.6	2.2	1.4	2.2	1.2
4	4.0	4.0	3.0	4.1	4.7	4.3	3.0	2.8	2.8	3.3	2.7	2.3
5	1.0	1.5	.8	1.4	1.6	1.7	.3	1.0	.8	1.0	.9	.5
6	1.0	1.4	2.9	2.4	3.9	4.1	3.7	3.4	1.7	2.3	1.8	2.1
7	3.0	2.7	4.2	4.0	4.9	9.9	3.7	3.2	3.7	2.9	2.7	3.2
Vowel Categories												
1	5.0	6.3	3.4	6.0	5.2	4.9	4.0	4.8	5.8	4.6	4.3	5.4
2	0.0	1.0	.3	.9	.8	1.1	1.0	1.8	.6	.8	.3	.9
3	0.0	2.1	.9	1.0	.5	.7	1.0	.4	.3	.2	.5	.2
4	3.0	1.0	2.3	1.5	1.9	1.7	1.3	.5	.5	.5	.8	.7
5	1.0	1.0	.1	.3	1.0	.9	1.0	.1	.5	.8	1.5	.4
Average intelligibility	23	29	48	37	28	23	18	30	23	27	33	22

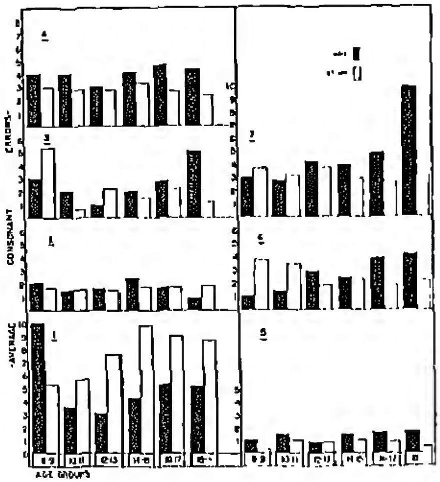

FIGURE 7

Average Number of Consonant Errors Per Pupil Arranged According to (*a*) Two-Year Age Groups; and (*b*) Error Categories

The tendency of consonant errors to increase with age (Figure 6) shows some modification when errors are broken down into categories.

Categories:

1. Surd-sonant	1. Substitution
2. Substitution	2. Diphthongs
3. Nasality	3. Diphthongization
4. Compound consonants	4. Neutralization
5. Abutting consonants	5. Nasality
6. Arresting consonants	
7. Releasing consonants	

throughout the age range except at the 8-9-year level. The maximum number of errors are made by the 14-15-year age-group at Mt. Airy, and by the 16-17-year age-group at Clarke.

Consonant substitution (Category 2, Figure 7) remains at a constant level throughout the entire age range for Mt. Airy pupils, but drops to a lower frequency with the older Clarke pupils.

The trend of nasality (Category 3, Figure 7) is reversed in the two school groups. There is an increase with age at Clarke and a decrease at Mt. Airy.

Errors involving compound consonants (Category 4, Figure 7) show minor variations in both school groups. Clarke pupils show

an increase in frequency up to the 16-17-year level, while at Mt. Airy the frequency decreases following the 14-15-year level.

Errors involving abutting consonants (Category 5, Figure 7) remain relatively constant for both school groups.

Errors involving the dropping of the arresting consonant (Category 6, Figure 7) show reversed trends in the two school groups. The frequency increases with age at Clarke and decreases at Mt. Airy.

Category 7, non-function of the releasing consonant, is similar in trend to Category 6 above. The increase in frequency of errors with age is more sharply defined among the Clarke pupils.

To summarize: The frequency of consonant errors appears to increase with age among the pupils of both schools studied. This tendency is marked among the Clarke pupils. Likewise, speech intelligibility, which is inversely related to frequency of errors, tends to decrease with age. Analysis of the relationship of the individual error categories and age show that for Clarke pupils all error categories increase in frequency with the exception of the 2nd (consonant substitution) and the 5th (malarticulation of abutting consonants). Category 1 (surd-sonant distinction) was the only one which showed a definite increase with age among the Mt. Airy pupils, while Categories 6 and 7 (dropping the releasing and arresting consonants) show a reduction in frequency with age.

This general summary of the data, showing an increase in consonant errors with age and the consequent lowering of speech intelligibility especially among Clarke pupils, makes it appear that the speech of deaf pupils deteriorates with age, even while the pupils are still in school. This fact, however, is not clearly established from the above analysis. Two factors other than age are possibly responsible for this *apparent* deterioration, namely: (*a*) The unequal distribution of pupils with different degrees of hearing-loss within the individual age-groups; and (*b*) the variability of the degree of speech intelligibility among individuals making up these sub-groups. A preponderance of profoundly deaf pupils in any age group, or conversely a relatively greater number of hard-of-hearing and partially deaf pupils in any age group would lower or raise the average at that level with reference to the general level. Furthermore, it is possible that among the older pupils studied, especially the profoundly deaf, some were poor speakers not because of any deterioration with age but because they have been poor speakers

TABLE 7

Showing the Relation of Consonant Errors and Speech Intelligibility to Age and Degree of Hearing Loss

The number of pupils at each age level is subdivided according to three degrees of deafness. Average intelligibility score per pupil, average number of consonant errors per pupil, and total averages are shown.

Hearing loss	A			B			C			Total		
Age group	No. pupils	Intell.	Cons. errors	No. pupils	Intell.	Cons. errors	No. pupils	Intell.	Cons. errors	No. Pupils	Intell.	Cons. errors
Clarke												
8- 9	—	—	—	—	—	—	1	23	24	1	23	24
10-11	—	—	—	1	52	12	7	26	17	8	29	17
12-13	3	80	11	6	50	16	7	34	20	16	48	17
14-15	4	75	7	6	45	18	14	23	27	24	37	21
16-17	2	76	9	6	40	27	15	17	26	23	28	25
18-	2	83	6	1	28	53	12	12	33	15	23	31
Totals	11	78	8	20	44	22	56	21	25	87	33 *SD* 26	23 *SD* 12
Mt. Airy												
8- 9	—	—	—	2	28	14	4	13	27	6	18	23
10-11	1	42	11	8	52	10	9	8	33	18	29	18
12-13	1	80	8	6	52	18	10	11	23	17	23	20
14-15	5	62	13	7	35	20	8	6	28	18	27	23
16-17	5	79	3	7	47	18	12	14	27	22	33	21
18-	2	74	1	6	52	15	16	5	24	24	22	20
Totals	10	69	7	36	43	16	59	9	25	105	26 *SD* 28	21 *SD* 10

from the very beginning. It is possible to analyze the data for the effects of these two factors and to determine whether or not there is a *real* deterioration of speech with age.

The effects of the first factor upon the distribution of average shown in Figure 6 can be determined from the data themselves. Table 7 shows the number of pupils, the average consonant errors and the average intelligibility scores for individuals within the several age groups according to the three degrees of hearing loss. The number of pupils falling into these three divisions vary widely among the different age groups. The number of profoundly deaf pupils (Group *C*) is relatively stable throughout the age range from the 10-11-year group upward, ranging from 7 to 16 pupils. The number of hard-of-hearing pupils (Group *A*) ranges from 0 to 4, while the number of partially deaf pupils (Group *B*) ranges from 0 to 8. Taking Groups *A* and *B* together the number ranges from 0 to 10 for the different age levels. This distribution is such, therefore, that the ratios of Groups *A* and *B* to Groups *C* are by no means constant at the several age levels.

The unequal distribution, as it appears in age-groups 10-11 (Mt. Airy), in 12-13 and 14-15 (Clarke and Mt. Airy), and finally in the 16-17-year group (Mt. Airy), is such that it tends to raise the general level of intelligibility and to lower the average consonant errors per pupil. On the other hand, the preponderance of profoundly deaf children in the 16-17, and in the 18-year age-groups, especially among the Clarke pupils, would tend to lower the intelligibility averages and raise the average consonant errors for each of these groups.

Because of this unequal distribution it is necessary to determine the trends of speech ability and age separately for each of the three degrees of hearing-loss. These data have been tabulated (Table 7) and plotted graphically in Figure 8. The averages for the most part are based upon a small number of pupils, hence the graphs show trends rather than statistically reliable differences.

Groups *B* and *C*, Clarke, show an increase in consonant errors with age and a tendency for the speech of the older pupils to become less intelligible (Figure 8). Group *A* reverses this trend, showing a slight improvement in speech with age.

Mt. Airy pupils show a slight decrease in consonant errors with age in Groups *A* and *C*, while Group *B* shows no definite trend (Figure 8). There is no apparent increase in consonant errors

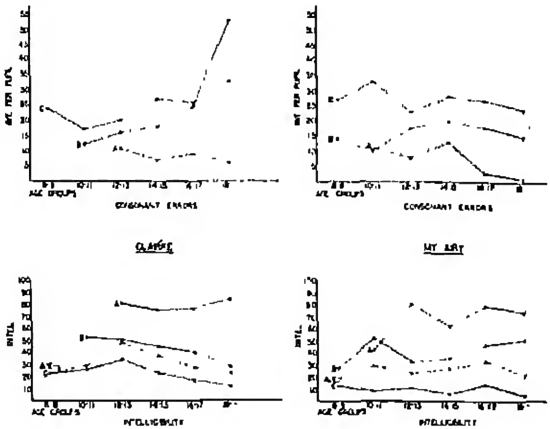

FIGURE 8

CHANGES IN ARTICULATORY ERRORS AND SPEECH INTELLIGIBILITY ACCORDING TO AGE AND DEGREE OF HEARING LOSS: THE DATA FOR THE TWO SCHOOLS ARE PRESENTED SEPARATELY

Consonant errors: Data plotted separately for groups with three degrees of hearing loss. (See Figure 6 for average consonant errors and age.) As consonant errors increase intelligibility scores decrease.

Intelligibility: Data plotted separately for groups with three degrees of hearing loss. The broken line represents averages of Groups *A*, *B* and *C*.

in either of the three divisions. Intelligibility scores, likewise, do not decrease with age among the Mt. Airy pupils. On the contrary, there is a tendency toward more intelligible speech among the older pupils of Groups *A* and *B*, while Group *C* remains relatively constant at a rather low level of intelligibility.

The evidence presented in Figure 8 indicates a deterioration in speech with age at Clarke among the partially deaf and profoundly deaf pupils. There is an indication of improvement in speech with age among the hard-of-hearing and partially deaf pupils at Mt. Airy, while the speech of the profoundly deaf remains constant. The fact that pupils in Groups *A* and *B* at Mt. Airy enjoyed, at the time of this study, a wider and a more systematic use of hearing aids than these same groups at Clarke in all of their classroom work probably accounts for the differences in the trends of speech intelligibility with age. The use of hearing aids, or the lack of it, cannot account, however, for the trends in speech intelligibility and

consonant errors noted among the profoundly deaf pupils of both schools.

The second factor mentioned above, the variability of speech proficiency, or intelligibility, among the individual pupils within the profoundly deaf groups, may possibly account for the apparent deterioration in the speech of the profoundly deaf pupils. It may be argued that an excessive number of poor speakers among the older pupils at Clarke does not mean that the speech of these pupils has deteriorated with age; it may be argued, rather, that they have been poor speakers all through their school years.

Obviously, a test of the type described in this paper would not reveal data concerning the previous status of the speech of any of the pupils. Periodic speech tests covering a period of years would be essential in order to show changes in the speech development of individuals. Data which partially meet this requirement are available for the Clarke pupils as a result of the annual speech-testing program initiated in 1935 and continued each year up to the present.[4]

A comparison of the test scores of a group of pupils over the six-year period (1935-1940) will determine whether or not there is a general tendency toward speech deterioration with age. Test scores are not available for Mt. Airy pupils.

Test scores are available over this period for Clarke pupils listed in the 10-11, and 12-13 and for some members of the 14-15-year age-

---

[4]The annual intelligibility tests are conducted as follows: "Each pupil reads 10 unrelated sentences to four auditors, students in the Teacher Education Department. Two of the auditors look at the pupil while two merely listen during the tests. Each sentence is repeated three times by the pupil and the auditors write down what they hear, or understand, at each reading."

"The scoring of the intelligibility tests is as follows: Sentences understood on the first reading are given a score of 10; those understood on the second reading receive a score of 5; and those not understood until the final reading are given a score of 2. No score is given unless the full sentence is understood. Finally, an average score taken from the four audition-records is determined for each child" (1, p. 35).

The validity of the intelligibility tests is determined by correlating test scores with teachers' rankings of individual pupils on the basis of speech intelligibility. A rank correlation between .70 and .80 is usually obtained.

The reliability of the intelligibility tests is determined by correlating test scores obtained from the same group of pupils in two consecutive years. The test materials and methods used are constant from year to year. A different group of auditors constitute the only change in the testing procedure. The correlation between test scores of 1939 and 1940 for a group of 96 pupils is .84, *PE* .02.

groups. Comparison of the speech intelligibility scores can be made for a total of 34 pupils; 19 of these are classified as profoundly deaf, 10 as partially deaf, and 5 as hard-of-hearing. The 1940 average intelligibility scores for this group of 34 pupils show no decline over the 1935 scores. Scores of some individual pupils, however, show gains, while those of others show losses. Of the 19 profoundly deaf pupils 12 showed gains ranging from 2 to 19 percentage points. The average gain for this group is 12 percentage points. The remaining profoundly deaf pupils showed losses ranging from 3 to 30 percentage points. The average loss for this group is 17 percentage points. The 1940 tests for the entire group of 19 profoundly deaf pupils showed a gain of 3 percentage points over the 1935 scores.

Of the 10 partially deaf pupils, for whom scores were available for comparison, four showed losses ranging from 7 to 18 percentage points, while six showed gains ranging from 5 to 14 points. Four of the five hard-of-hearing pupils showed gains ranging from 1 to 7 points during this six-year period, while one showed a loss of 9 percentage points.

It is evident from the above analysis that speech intelligibility, as it is tested by this method, may increase, decrease, or remain static as pupils continue in school over a six-year period. No general trend for the group of pupils as a whole is apparent. Wide individual differences of both the degree of speech intelligibility at any given time, and the amount of individual progress or retardation in speech intelligibility over a period of years, appear to be of far greater significance for study and analysis than general group trends. The degree of speech proficiency existing within a group of pupils at any given stage in their speech development, as well as the amount of progress that will be made over a period of years will depend upon factors which produce these individual differences within the group. These factors are not clarified in the above analysis, and a more detailed investigation will be necessary for their clarification. We may suggest, however, that factors, such as the intelligence of the pupil, the amount of residual hearing available for use, the degree to which this residual hearing is used in speech development, the relative amount of time spent upon speech training as the pupil advances from the lower to the higher grades, and the extent to which speech is used as a means of daily communication, all of these factors, determine the trends of speech intelligibility with age.

In discussing changes which occur in the speech of deaf children with age, Haycock (8 pp. 255-261) says, (p. 255): "It is commonly observed that the speech of deaf-born pupils become less intelligible as they advance from the lower to the higher classes in school; and this increasing difficulty to understand their speech is ascribed to its deterioration." In attempting to explain this statement Haycock suggests that the less intelligible speech of the older pupils is not due to a process of *deterioration*, but rather to factors inherent in the progress of deaf children as they advance from the lower to the higher grades. These factors Haycock lists as: differences between the range of vocabularies, grammatical structures, degree of phonetic complexity, and the relative amount of time spent purely on speech instruction in the lower and the more advanced grades. He believes these factors account for the differences in speech intelligibility among the younger and older pupils. He considers that "inadequate and insufficient training and preparation to meet a continually growing and expanding set of varied speech requirements" are responsible for the apparent deterioration in the speech of the older pupils.

G. Analysis of Errors of Rhythm

An analysis of the errors in the speech of deaf pupils would be quite incomplete if errors of rhythm were omitted. One of the serious defects in the speech of deaf children is the labored, malphrased, and badly accented sentences. The apparent disregard for the normal rhythmic patterns helps make their speech unintelligible. It was a relatively simple matter in making an analysis of the speech records of the subjects in this survey to grade each sentence for rhythmic errors. The rhythmic patterns, or the lack of them, stand out clearly as the records are played back several times; and it becomes an easy matter to transcribe the actual rhythms in terms of commonly used symbols. Each of the test sentences, some 1,900 in all, were transcribed in this manner and classified according to three rhythm categories. A brief discussion of the nature of speech rhythm will clarify the description of the rhythm categories.

Obviously rhythm is not primarily a matter of sound since congenitally deaf children learn to execute rhythms of various kinds such as musical rhythms, dances, calisthenic drills and even recite verse with a definite feeling for the rhythmic patterns. Their ordinary speech is often rhythmic although it may not follow the

normal English rhythmic patterns. The fact that they do not learn the English speech rhythms easily or as quickly as normal children does not mean that speech rhythm is dependent upon sound. It means rather that deafness blocks the normal avenues by which speech rhythm is learned; it means also that teachers of the deaf have not yet developed wholly adequate methods of teaching speech rhythm.

Sounds can arouse a rhythm and can be grouped into rhythmic sequences, but rhythm has a far wider field than that of sound alone, and is essentially a matter of the grouping of movements (16, 14). Therefore, it seems to us there is nothing about the rhythms of speech that the deaf child cannot learn.

Rhythm is basically the grouping of movements about a main component which is said to be accented. The accented movement is the stronger, or the stressed impulse, and marks the regular repetition of "feet" in verse. In prose the "feet" are not repeated in any regular recurrent pattern. The terms "accentuation" and "subordination" describe the dynamic processes involved in rhythmic grouping (14, p. 293). By grouping is meant that two or more unit impulses are joined together forming a larger unit, commonly called a "foot," in which the lesser component elements, or movements are subordinated to the stronger, accented component. The expressions: *on the first, in school, take it, aptitude*, spoken normally with accents as indicated, are examples of unit groups or feet illustrating the phenomenon of grouping.

Rhythmic feet are, in turn, grouped into larger units called "phrases" or "breath-groups" in speech. The smaller unitary movements in speech are the pulsations for the syllables (17, p. 29; 18, p. 246). The slower movement of expiration binds the syllables which are grouped into rhythmic feet into the larger unitary movement of the phrase. The sentence: *The boy took his brother to the movies*, spoken on a single expiratory movement with the syllables grouped into feet as indicated, becomes a breath-group in which the individual feet are closely joined by the unitary movement of expiration. This slow movement of breathing-out gives the phrase its unity.

In English speech the word-accent is definitely fixed so that the accent of the word is observed in the phrase although the word may be lost in the breath-grouping. Strictly speaking, the syllable which carries the word-accent continues to be accented as the word be-

comes a part of the phrase. The word accents of *bróthĕr* and *móvĭes* in the example given above are retained and help mold the rhythm of the phrase in which they appear. A shift of these accents to the second syllables, as in *brŏthér* and *mŏviés* would not only modify the words themselves but would violate the normal English rhythm of the phrase. Foreigners learning to speak English have great difficulty with word-accent.

The rate at which movements recur in a sequence vitally affects the rhythmic grouping. There is a lower and an upper limit to the rate of succession at which the grouping of events in a sequence is possible (14, pp. 295-296). Movements, sounds, or other events occurring in a sequence as slowly as 20 to 40 per minute cannot be grouped rhythmically; at such a low rate each unitary event becomes an accented single unit-foot. Again, a rapid series of movements which recur at rates as high as 6 to 8 per second cannot be grouped into rhythmic units since the rapid rate does not permit time for an accented movement which requires a longer duration.

Stetson, 1928 (15, pp. 205-206), in discussing the influence of rhythm on the pronunciation of a language says:

> The rhythm is certainly one of the most fundamental characteristics of the pronunciation of a language, and is often most difficult for a foreigner to acquire. The play of the word accents, the rhythmic grouping of the phrases, the differences in the length of the syllables are all difficult, and all important for good "accent." And it is not the case that one can first master the "elements" of the pronunciation, the "sounds" and then set them in the rhythm. It is easy to see that rhythm has a vital influence on details of pronunciation; the word accent often determines the function of the consonant as arresting or releasing and also determines the syllable in which the consonant shall function; the rhythm at high speed determines the slurring or the full pronunciation of syllables.

The question arises as to whether the "Elements Method" of speech teaching, common in most schools for the deaf, is in accord with the basic principles of the development of speech rhythm as expressed in the above quotation, namely, *elements of pronunciation, the "sounds"* are not mastered first then *set in the rhythm*. The Elements Method stresses the teaching of individual elements; the elements are then combined to form words, finally words are combined to form phrases. Teachers insist upon "accuracy" for these

elements by which they mean that each sound must have a relatively fixed durational, qualitative, and intensity value before they are combined into syllables and words. This procedure overlooks the very important and basic fact that *syllables rather than "sounds" carry the rhythm of speech,* and that syllables are strong or weak, long or short, depending upon the dynamic patterns of rhythmic grouping. For instance, the sentences: *My father is taller than I am,* and *The kitten can run fast,* would be spoken normally with the rhythmic grouping as indicated. The division of the sentences into unit-groups (feet), while somewhat arbitrary, may be also as indicated. The accented syllables will not only be uttered more forcibly, but they will also have a greater duration than the unaccented syllables. The individual sounds, the consonants and the vowels, are likewise long or short, stressed or unstressed, depending upon their occurrence in accented or unaccented syllables. For instance the vowel *i-e,* the pronoun *I* (first sentence), is of much greater duration than the same vowel in the pronoun *my* when the sentence is spoken as indicated. The *k* in *kitten* (second sentence) likewise is more vigorous and has a longer duration than the same sound in the word *can* in the same sentence.

A further illustration of the effects of the rhythm of the phrase upon individual sounds is offered by the phrase *I want a new hat.* When the verb *want* is accented in this phrase the arresting compound consonant *nt* is retained in that position. A shift of the accent to *new,* however, leaving the verb unstressed, normally causes the *t* of the combination to shift over to the following syllable and the phrase becomes *I wan ta new hat.* Again, a consonant may be dropped entirely by a slight change of accent as in the phrase: *I want to go home* changed to: *I want to go home.* The retention of the accent on *want* permits sufficient time for the retention of the final *t* which doubles with the initial *t* of *to,* the following syllable. When the accent is shifted, however, and placed upon the verb *go* the rate of the unaccented syllable *want* and *to* is much too rapid for the double consonant and the phrase becomes: *I wan to go home.*[5]

Individual sounds must accommodate themselves to the demands of the rhythmic grouping of the syllables in which they occur.

[5]A detailed discussion of the effects of word-accent upon syllables and its relation to individual phonemes has been presented by Stetson (15, pp. 140-205).

They may be lengthened, shortened, shifted from one syllable to another, or even dropped out of the coördination altogether. Speech rhythm must suffer, therefore, when the speaker by force of habit insists upon giving each individual sound a definite stereotyped value. It is worth repeating that *syllables rather than individual sounds carry the rhythm of speech.* Even normal word accent becomes distorted under conditions in which the speaker "punches out" each sound in a manner similar to that in which individual letters are "punched out" in typewriting. When words such as *father, morning, around,* etc., are spoken in a manner such that the sounds in both syllables are of equal duration and intensity the effect is that of accenting both syllables, or of shifting the normal accent to the unaccented syllable; the words lose both their acoustic and rhythmic form.

In this connection it is of interest to note that Dr. Bell, who advocated the elements method of speech teaching, was not only aware of the importance of rhythm, but pointed out the reasons for the lack of it in the speech of congenitally deaf children. In discussing the matter (2, pp. 14), he says:

> When I first entered upon the work of articulation teaching, I was very proud of the pronunciation of some of my congenitally deaf pupils. They had been drilled upon the elements and were able to pronounce words and sentences written in Visible Speech with absolute correctness, slowly, it is true, but with perfect elementary sounds. To my great mortification, however, I found that visitors generally preferred the imperfect gabble of some semi-mute to the elocutionary speech I had labored to impart.

Again, Bell says, p. 15:

> Ordinary people who know nothing of phonetics or elocution have difficulty in understanding slow speech composed of *perfect elemenatry sounds,* while they have no difficulty in understanding an imperfect gabble if only the *accent and rhythm* are natural (italics ours).

The point inferred here is that "perfect elementary sounds" from a phonetic or elocutionary point of view are incompatible with the dynamic processes of grouping, accentuation, and subordination of syllables. In the passages which follow the above quotations Bell discusses the wasted labor "bestowed upon unaccented syllables."

From the standpoint of rhythm the sentences spoken by the 192 deaf pupils studied in this investigation can be divided into three types of categories: (*a*) Those spoken with normal rhythm, e.g., *Ĭt wĭll nót ráin tŏdáy*; (*b*) those spoken with abnormal rhythm, e.g., *Ítŭ wíllŭ nótŭ ráin tódáy*; and (*c*) those spoken non-rhythmically, e.g., *Ít wíll nót ráin tódáy*.

Sentences spoken with *normal rhythm* included all those in which the accents were properly placed, in which the normal English grouping of syllables obtained, and in which the rate of syllable utterance was such that grouping was possible. Of 1,868 sentences available for rhythmic analysis 45 per cent were spoken with normal rhythm. This 45 per cent, or 837 sentences, accounted for 74 per cent of all sentences correctly understood by the auditors.

The rhythm of a sentence was classified as *abnormal* when the rhythmic pattern conflicted with, or did not conform to a normal English pattern for that sentence. This abnormality took one of four forms: (*a*) Sentences were broken up into short, unusual breath-groups; (*b*) word accents were misplaced, or normally unaccented syllables were accented; (*c*) adventitious syllables were added to words containing compound consonants, or between syllables connected by abutting consonants. The phrase *in the snow,* spoken as *inu the sunow* is an example. Often forms (*b*) and (*c*) were combined. (*d*) Syllables were omitted from polysyllabic words thus breaking up the normal rhythmic pattern. Of the sentences analyzed 36 per cent, or 668 sentences, were spoken with abnormal rhythm. This 36 per cent accounted for only 17 per cent of all sentences correctly understood by the auditors.

Sentences were classified as *non-rhythmic* when there was a complete absence of grouping, when each syllable in the phrase was spoken slowly and with the same degree of stress, and as a single breath-group. This type of utterance is either too slow for any semblance of rhythmic grouping or the "punch rhythm" with complete lack of variation of accent gives it a form utterly lacking in rhythmic characteristics. It may be characterized as *metronomic* speech. Nineteen per cent, or 363 sentences, of the total group were classified as non-rhythmic. These accounted for only 9 per cent of the sentences correctly understood by the auditors.

To summarize: Sentences spoken rhythmically correct by deaf pupils have almost a four to one (3.5 to 1) advantage of being understood over those spoken with incorrect rhythm. Slightly

less than half (45 per cent) of all sentences spoken by 192 deaf pupils were spoken with normal rhythm, yet these accounted for three-fourths (74 per cent) of all the sentences understood by the auditors. The remaining sentences spoken with abnormal rhythm or with no rhythm (55 per cent) accounted for only 26 per cent of all sentences understood by the auditors.

A summary of the data for the two school groups taken separately is given in Table 8. There is a striking similarity in the speech

TABLE 8

NUMBER AND PER CENT OF SENTENCES SPOKEN WITH THREE TYPES OF RHYTHM, AND PER CENT OF SENTENCES WHICH WERE UNDERSTOOD BY THE AUDITORS WHICH FALL INTO THE THREE RHYTHM CATEGORIES

Rhythm category	No. of sentences spoken	% of total	% of those understood
	Clarke		
1. Normal	385	46	73
2. Abnormal	329	39	20
3. Non-rhythm	127	15	7
Total	841	100	100
	Mt. Airy		
1. Normal	452	44	75
2. Abnormal	339	33	14
3. Non-rhythm	236	23	11
Total	1,027	100	100

rhythm data of the pupils of the two schools in this general analysis.

Defective rhythm, of course, was not the only cause for the failure of the auditors to understand the speech samples. Indeed, not all of the sentences spoken with correct rhythm were understood by the auditors. Articulatory errors, previously discussed, were present in most of the sentences. In fact, those sentences free from defects of rhythm contained fewer articulatory errors, and sentences spoken with abnormal rhythm or non-rhythmically contained a greater number of articulatory errors. The presence of articulatory errors along with errors of rhythm would natually condition the intelligibility of the sentence to a degree which cannot be attributed wholly to its rhythmic pattern. In order to determine the true extent of the relationship between articulatory errors and errors of rhythm correlation coefficients between these two variables were calculated.

TABLE C

Clarke: *Consonant errors vs. correct rhythm*	—.41, *PE* .06
Vowel errors vs. correct rhythm	—.60, *PE* .05
Mt. Airy: *Consonant errors vs. correct rhythm*	—.69, *PE* .03
Vowel errors vs. correct rhythm	—.53, *PE* .05

The coefficients were significant in each case (Table *C*). The negative coefficients of correlation, although not extremely high, are reliable, and indicate that sentences spoken with correct rhythm are likely to have fewer articulatory errors than those spoken with abnormal rhythm.

Both articulatory errors and errors of rhythm, therefore, are operative in determining the degree of intelligibility of each sentence. It would be interesting to know just what are the relative contributions to speech intelligibility of these two variables. The methods of partial correlation offered a means of studying this question with regards to the present data. By statistically ruling out one or more of the variables it was possible to show the true correlation between intelligibility and consonant errors, intelligibility and vowel errors, and intelligibility and correct rhythm. Partial coefficients of correlation were calculated from the data of the two schools separately. Table 9 shows the intercorrelation between intelligibility, rhythm, consonant errors, and vowel errors; and finally partial correlation coefficients between intelligibility and each of the three variables with the remaining two ruled out, or held constant.

The data in Table 9 indicates that rhythm is a rather important item in the speech intelligibility of deaf pupils. The intercorrelation of rhythm and intelligibility is .71 and .76 for Clarke and Mt. Airy respectively. When the effects of consonant and vowel errors are ruled out (partial correlation method) there is still a relatively high degree of correlation betwen the two variables. The actual figures are .56 (Clarke) and .52 (Mt. Airy). These figures show that the relationship between speech rhythm and speech intelligibility is just as great as that between consonants and intelligibility, and considerably greater than that between vowels and intelligibility.

It has been noted by many observers, among them Bell (2), and Story (19), both of whom had had considerable experience in teaching speech to deaf children, that consonants are more important than vowels in determining speech intelligibility. The data of this investigation bear out this observation and furnish quantitative evi-

TABLE 9

Correlations

	Clarke 1 Intelligibility	Clarke 2 Rhythm	Clarke 3 Consonants	Clarke 4 Vowels		Mt. Airy 1 Intelligibility	Mt. Airy 2 Rhythm	Mt. Airy 3 Consonants	Mt. Airy 4 Vowels
Intercorrelations									
1. Intelligibility	—	.71	—.70	—.61	1. Intelligibility	—	.76	—.71	—.56
PE		.04	.04	.05	*PE*		.03	.03	.05
2. Rhythm	.71	—	—.41	—.60	2. Rhythm	.76	—	—.69	—.53
PE	.04		.06	.05	*PE*	.03		.03	.05
3. Consonants	—.70	—.41	—	.47	3. Consonants	—.71	—.69	—	.68
PE	.04	.06		.06	*PE*	.03	.03		.04
4. Vowels	—.61	—.60	.47	—	4. Vowels	—.56	—.53	.68	—
PE	.05	.05	.06		*PE*	.05	.05	.04	
Partial correlations									
	$r_{12.34}$	.56				$r_{12.34}$	.52		
	$r_{13.24}$	—.60				$r_{13.24}$	—.30		
	$r_{14.23}$	—.20				$r_{14.23}$	—.12		

dence in its support. When the effects of rhythm and consonants are ruled out the partial correlation between vowel errors and intelligibility is rather low: —.20 for Clarke pupils and —.12 for Mt. Airy pupils. Similar coefficients for consonant errors and intelligibility, on the other hand, are: —.60 (Clarke) and —.30 (Mt. Airy). If we may accept these partial correlation coefficients as indicative of the relative importance of consonants and vowels in the interpretation of speech it is clear that the former are of much greater importance.

Speech rhythm, as shown by the data, becomes an important factor in speech intelligibility. The data presented above applies, of course, to the pupils and speech materials of this study. Several questions, however, are raised: Can we say the results are typical of the speech of deaf children in general? What are the practical implications for speech teaching and speech correction for deaf children?

In answering these questions some degree of speculation is involved. It is safe to say, however, that the speech of the pupils studied is representative of speech teaching in America at the present time. Both schools have long traditions in teaching speech. Their teachers and administrators have been pioneers in the education of the deaf. The pupils studied were representative of the student populations for the years 1936 and 1937. The test materials used were simple; they were worked out in collaboration with teachers who knew the type of language forms and vocabularies familiar to their pupils.

As to the practical application of these findings regarding speech rhythm there are clear indications that greater attention and stress upon normal speech rhythm should greatly improve the speech of deaf children. The truth of this is eloquently suggested by the fact that while only 45 per cent of all sentences were spoken with normal rhythm, this 45 per cent accounted for 74 per cent of sentences understood (Table 8). The need for basic training in speech rhythm is also indicated by the data. This need is sharply defined in the rhythm data of the profoundly deaf pupils, Groups *C*. Of the sentences spoken by these pupils only 30 and 23 per cent were spoken with normal rhythm by Clarke and Mt. Airy pupils respectively. These, in turn, accounted for 52 and 41 per cent of all sentences understood by the auditors.

Greater attention to the development of normal speech rhythm would mean, among other things, a better application of the funda-

mentals of speech rhythm. It would mean that less emphasis be placed upon the teaching of single sounds, "elements," and greater emphasis upon teaching the dynamic functions of consonants and vowels as they combine in the chain of syllables which make up breath groups. *Individual "sounds" are not mastered first and then "set in the rhythm" as current "analytic" methods of speech teaching implies.* On the contrary, syllables, the primary elements of speech carry the rhythm and the "elements" or sounds must accommodate themselves to the rhythmic demands of the chain of syllables in the phrase. Individual "sounds" are shortened, lengthened, caused to fuse with *other sounds, shifted from one syllable to the other and* even forced to drop out completely by the varying rhythmic patterns and varying rates of normal speech. These shifts and changes become impossible *when the speaker attempts to give stereotyped values to each individual sound.*

Speech rhythm is a specific form of rhythm; it is based upon the syllable rate, the word accent, and the proper grouping of syllables about this accent in the formation of breath groups regulated by the breathing muscles. "Rhythm Classes" in which deaf children *learn complicated musical rhythms for brief periods each day probably* are of little value as such since the patterns learned are not speech rhythms and can have little or no application in speech development. *Since speech rhythms are produced and controlled by the* breathing muscles, the training of these muscles in the *proper forms of breath control and phrasing* becomes an *important part of speech training.* Furthermore, the articulatory organs which produce the sounds should not be trained in isolation in the "development" of the individual sounds since these sounds do *not occur in isolation in speech. An important part of speech* training consists in coördinating the two basic movement series of speech, the syllable movements and the articulatory movements, into a *functional whole which we call the* speech mechanism. The teaching of speech rhythm, therefore, cannot be actually separated from the teaching of speech. From the very beginning of speech development *speech rhythm, that is, accentuation, grouping* and phrasing of syllables, should be a vital and inseparable part of speech training, rather than an independent project at which classes spend 15 to 30 minutes each day in the "*rhythm room.*"

1. *Speech Rhythm and Degree of Hearing Loss*

Analysis of the rhythm data shows that the development of normal speech rhythm, like the other aspects of speech previously discussed, is affected by the amount of residual hearing. One method of showing this relationship is that of the degree of correlation between the number of sentences (in 10) spoken rhythmically correct and the amount of residual hearing as expressed by the audiometric scores. Correlation coefficients of .61 and .69 were found for these two variables for Clarke and Mt. Airy pupils respectively. This indicates that the presence of even a small amount of residual hearing is an asset in the development of normal speech rhythm.

A detailed analysis of the effect of hearing loss upon speech rhythm is presented in Table 10. The average number of sentences (in 10) spoken with normal rhythm, and the number of correct auditions for each sentence (based upon the records of five auditors) are presented for each of the three divisions of hearing loss, Groups *A*, *B*, and *C*. The figures for the two school groups are very similar, showing similar trends, and with a slight advantage in favor of Clarke pupils for Groups *A* and *C*. Group *B*, Mt. Airy, is slightly superior to the corresponding Clarke Group. The number of sentences per pupil spoken with normal rhythm in Groups *A*, *B*, and *C* stands in the ratio of 3:2:1.

The effect of rhythm upon the intelligibility is demonstrated again in the figures indicating the number of correct auditions per sentence. The effect of abnormal or non-rhythmical speech is less apparent in sentences spoken by hard-of-hearing pupils (Groups *A*) because of the few articulatory errors in these sentences. On the other hand, sentences spoken with normal rhythm by profoundly deaf pupils (Groups *C*) were two to three times more intelligible than those spoken with abnormal or with no rhythm.

The data show clearly that even a small amount of residual hearing enables "deaf" children to acquire the normal rhythmic patterns of speech more completely than those deprived of this avenue. This fact, however, is not to be interpreted as meaning that profoundly deaf children cannot learn the normal rhythmic patterns of speech. Indeed, they do learn rhythmic patterns in verse, calisthenic drills, dances, and even musical rhythms. The rhythms of speech are probably not more difficult than some of these. Proficiency in the execution of the rhythms of dancing or music, however, will be of

TABLE 10

ANALYSIS OF SPEECH RHYTHM ACCORDING TO DEGREE OF DEAFNESS

Degree of hearing loss		(1) Normal rhythm			(2) Abnormal rhythm			(3) Non rhythmic		
		No. sents.** spoken	Av. per pupil	Av. correct auditions* per sent.	No. sents.** spoken	Av. per pupil	Av. correct auditions* per sent.	No. sents.** spoken	Av. per pupil	Av. correct auditions* per sent.
					Clarke					
Group *A*	11	95	8.6	3.9	10	.9	3.8	0	0.0	0.0
Group *B*	20	124	6.2	2.3	54	2.7	1.1	18	.9	1.9
Group *C*	56	166	3.0	1.7	265	4.7	.7	109	1.9	.6
Total	87	385	4.4	2.4	329	3.8	.85	127	1.5	.82
					Mt. Airy					
Group *A*	10	87	8.7	3.8	6	.6	2.0	6	.6	2.1
Group *B*	36	230	6.4	2.7	81	2.3	1.1	43	1.2	1.6
Group *C*	59	135	2.3	.9	252	4.3	.4	187	3.2	.4
Total	105	452	4.3	2.3	339	3.2	.6	236	2.3	.6
Total both	192	837	4.4	2.4	668	3.5	.7	363	1.9	.7

*The number of correct auditions per *sentence* is based upon the average of five auditors for each sentence. This was necessary since originally there were an unequal number of auditors, usually from 5 to 10 for each child.

**Each pupil read ten sentences.

little help in learning speech rhythms since there is so very little direct relationship between them. The effects of exercising the legs or the arms in learning a movement sequence involved in a complicated rhythmic pattern are not transferred to the speech mechanism. Even if such transfer were possible it would be of little use since speech rhythms are quite different in dynamic pattern. The way for pupils to learn speech rhythms is to have them practice speech rhythms *specifically, to practice from the beginning of speech training the* proper grouping of syllables into breath groups. Individual "elements" embodied in syllables, while not to be ignored, are subordinate to the rhythmic grouping of syllables and undergo changes of duration and function, changes dictated by the demands of rhythm. Hearing aids are especially helpful in teaching speech rhythm. Even the profoundly deaf can get the rhythm of a phrase though none of the sounds are understood. The tactile, visual, and kinaesthetic senses are also available for presenting the rhythmic patterns of phrases.

2. *Analysis of Rhythm Data According to Age*

Table 11 presents an analysis of the data from the two school groups according to two-year age-groups. The average number of sentences (in 10) per pupil spoken with each type of rhythm and the average number of auditors (among 5) who understood each sentence are listed in separate columns. Mt. Airy pupils show slight increase in the number of sentences spoken with correct rhythm with age up through the 18-year group, but the number of correct auditions shows no definite trend with age. The 12-13-year age-group, among the Clarke pupils, show the highest number of rhythmically correct sentences and a slight decrease from this group on to the older pupils. Likewise the number of correct auditions per sentence is highest at the 12-13-year level and decreases among the older groups. There is a general decrease with age in the number of sentences spoken non-rhythmically in both school groups. This indicates that pupils are learning, as they grow older, to group syllables into a form of rhythmic pattern. These patterns, however, are not necessarily the normal English forms of grouping and accentuation. Accentuation and grouping of movements is a phenomenon which will naturally occur whenever a series of movements (syllables) are executed at rates at which grouping is possible. Increasing familiarity with articulatory movements, vocabularies, and

TABLE 11

ANALYSIS OF SPEECH RHYTHM DATA ACCORDING TO TWO-YEAR AGE-GROUPS

	(1) Normal rhythm				(2) Abnormal rhythm			(3) Non-rhythmical		
Age groups	No. pupils	No. sents. spoken	Av. per pupil	Av. correct auditions* per sent.	No. sents. spoken	Av. per pupil	Av. correct auditions* per sent.	No. sents. spoken	Av. per pupil	Av. correct auditions* per sent.
					Clarke					
8- 9	1	0	0.0	0.0	0	0.0	0.0	10	10.0	1.1
10-11	8	21	2.5	1.8	36	4.5	.9	22	2.5	.9
12-13	16	91	5.7	3.5	51	3.2	1.0	13	1.0	2.0
14-15	24	103	4.3	2.5	80	3.3	1.2	39	1.6	1.0
16-17	23	110	4.8	1.9	95	4.1	.7	21	.9	.4
18-	15	60	4.0	1.8	67	4.5	.6	22	1.5	.14
Total	87	385	4.4	2.4	329	3.8	.85	127	1.5	.82
					Mt. Airy					
8- 9	6	16	2.7	1.3	16	2.7	.8	28	4.7	.8
10-11	18	76	4.2	2.7	63	3.5	.6	38	2.1	.6
12-13	17	59	3.1	1.6	67	3.9	.5	39	2.3	1.3
14-15	18	78	4.3	2.4	58	3.2	.7	41	2.3	.4
16-17	22	96	4.4	3.0	59	2.7	.6	62	2.8	.6
18-	24	127	5.3	1.9	76	3.2	.5	28	1.2	.3
Total	105	452	4.3	2.3	339	3.2	.6	236	2.3	.6

*Number of correct auditions per sentence is based on the number of correct auditions for five auditors for each sentence. It was necessary to reduce the number of auditors to a common number since the actual number varied from time to time.

grammatical forms permits the older pupils to repeat syllables and words at a more rapid rate of utterance than that of the younger pupils. Various forms of accentuation and rhythmic groupings will naturally occur under these circumstances. It is not inevitable, however, that the patterns thus formed will be those of normal English groupings unless they are specifically taught. In general, therefore, it can be said that there is a tendency among the Clarke and Mt. Airy pupils to speak more rhythmically with increasing years in school, but the rhythmic patterns used, especially by Clarke

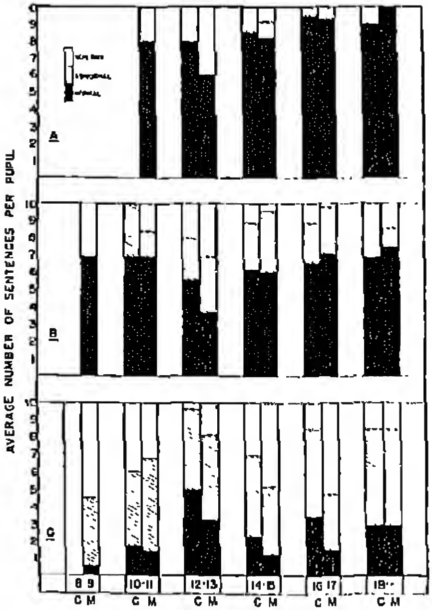

FIGURE 9

AVERAGE NUMBER OF SENTENCES PER PUPIL SPOKEN WITH THREE TYPES OF RHYTHM ARRANGED ACCORDING TO AGE GROUPS AND HEARING LOSS

C.—Clarke data.

M.— Mt. Airy data.

The number of sentences spoken with normal rhythm (Blackened area) increases with age; this number is reduced as deafness increases (Groups *A* to *C*).

The number of sentences spoken with abnormal rhythm (hatched area) changes very little with age, but decreases as the degree of deafness decreases.

The number of sentences spoken non-rhythmically (clear area) decreases with age and increases with degree of hearing loss.

pupils, are not more likely to be the normal or correct forms with increasing age. The increasing tendency toward the grouping of syllables into feet and breath groups with age is probably due to increasing familiarity with language forms, to an increasing fluency of speech, and to an increasing vocabulary with age. But there is little or no tendency for the older pupils to use the correct form of English accentuation and grouping more frequently than the younger ones.

The rhythm analyses for hearing loss and for age (Tables 10 and 11) are both presented graphically in Figure 9. The solid portions of the columns indicate the average number of sentences spoken with correct rhythm at each age level for the three degrees of hearing loss. The hatched portions indicate the average number of sentences spoken with abnormal rhythm. The clear portions indicate the average number of sentences spoken non-rhythmically. Two general tendencies are clearly shown in the figure: A decided increase in number of sentences spoken with normal rhythm with increasing amount of residual hearing; and a decrease in the number of sentences spoken non-rhythmically with increasing age, especially in Groups *B* and *C*. Group *A* shows most clearly an increase in normal rhythm with age.

H. Relative Difficulty of Individual Phonemes in the Speech of the Deaf

The data presented thus far deal with speech errors considered in terms of error categories. Comparisons of these categories have been made on the basis of their relative frequency, their relative importance for speech intelligibility, and between groups of pupils classified according to age and degree of hearing loss.

It should be of interest to make a further analysis in which individual phonemes are ranged on a *difficulty scale*. The position of any phoneme in the scale will be determined by its stability in the speech of deaf children. In other words, the number of times any phoneme is malarticulated or dropped in the speech of 192 deaf children, taking into account the relative frequency of occurrence of that phoneme, determines its place in the difficulty scale.

Zipf (25), in attempting to show a relationship between the frequency of occurrence and degree of complexity of phonemes met with difficulty in establishing a quantitative scale of "degree of complexity." He finally resorted (pp. 59-61) to a theoretical scale based upon what he called "a constellation of articulatory sub-

gestures" in which the number of separate articulatory events involved in the production of a phoneme determined its degree of complexity. On this basis he concluded that aspirated stops were more complex than unaspirated stops, *fortes* stops more complex than *lenes*, and voiced stops were more complex than voiceless stops. On the basis of this method of classification Zipf shows an inverse ratio between the frequency of occurrence of these sounds and their relative complexity in the several languages studied.

Since deaf children must learn to produce the phonemes largely or solely in terms of articulatory movements any differences as to degree of complexity of these movements should appear as the pupils use the sounds in speech. It becomes possible to work out a pragmatic *scale of difficulty*, therefore, in which each phoneme finds its place by virtue of its instability in the speech of deaf children as determined by the number of times each phoneme is malarticulated. Such a scale should have practical value for teachers of the deaf and at the same time be of theoretical interest to phoneticians, although a scale of difficulty for normal hearing children might be totally different.

By going over the data it was possible to determine the number of times each consonant or vowel was malarticulated by 192 deaf children. The several types of errors were ignored in this tabulation. Consonants which appeared as components of compound consonants were not included in the tabulations since such combinations are fusions of two or more consonants in which the true identity of the individual components is lost. Frequency tables showing the number of times each sound was dropped or mispronounced were made. Table 12 shows the consonants and vowels listed in the order of frequency of errors. The data obtained from the pupils of the two schools are listed first separately and then combined. There is a close agreement in the ranking between the two school groups in both consonants and vowels. The rank order correlation for consonants for the two schools is .88, *PE* .02 and for vowels it is .92, *PE* .014.

1. *Error Index Ranking of Phonemes*

The rankings of the data in Table 12, however, are based on frequency of errors of individual sounds and ignores the actual frequency of occurrence of these sounds in the sentences used as test materials. Such a ranking, therefore, does not give a true picture of the relative difficulty, or relative stability, of individual sounds in the speech of deaf children. A true ranking of each sound

TABLE 12

CONSONANTS AND VOWELS RANKED ACCORDING TO FREQUENCY OF ERRORS

Clarke			*Mt. Airy*			Combined		
			Consonants					
1.	l	165	1.	d	359	1.	d	502
2.	d	143	2.	l	275	2.	l	440
3.	h	140	3.	b	230	3.	b	335
4.	s	111	4.	h	159	4.	h	299
5.	b	105	5.	z	126	5.	s	235
6.	t	91	6.	s	124	6.	t	192
7.	th^2	86	7.	g	110	7.	z	182
8.	r	85	8.	t	101	8.	r	180
9.	n	75	9.	r	95	9.	g	171
10.5	sh	61	10.	y	55	10.	th^2	124
10.5	g	61	11.	w	49	11.	y	105
12.	k	57	12.	th^2	38	12.	n	102
13.	z	56	13.	sh	33	13.	sh	94
14.	m	55	14.	k	29	14.	w	88
15.	y	50	15.	m	28	15.	k	86
16.	f	41	16.5	n	27	16.	m	83
17.	w	39	16.5	j	27	17.	v	60
18.	v	38	18.	ch	26	18.	f	53
19.	p	30	19.	v	22	19.	ch	52
20.	ch	26	20.5	th^1	13	20.	j	50
21.	j	23	20.5	p	13	21.	p	43
22.	wh	14	22.	f	12	22.5	th^1	23
23.	–ng	12	23.	–ng	11	22.5	–ng	23
24.	th^1	10	24.	wh	7	24.	wh	21
25.	zh	0	25.	zh	0	25.	zh	0
			Vowels					
1.	i-e	121	1.	–i–	173	1.	–i–	249
2.	–i–	76	2.	i-e	118	2.	i-e	239
3.	a-e	71	3.	ee	83	3.	ee	151
4.	ee	68	4.	–u–	66	4.	–e–	114
5.	oo^1	53	5.	–e–	63	5.	a-e	112
6.	–e–	51	6.5	–a–	41	6.	–u–	104
7.	o-e	50	6.5	a-e	41	7.	oo^1	91
8.	–a–	45	8.	oo^1	37	8.	–a–	86
9.	–u–	38	9.	ur	35	9.	o-e	74
10.	aw	36	10.	–o–	34	10.	–o–	69
11.	–o–	35	11.	aw	29	11.	aw	64
12.	ou	26	12.	o-e	24	12.	ur	50
13.	oo^2	20	13.	ow	17	13.	ou	43
14.	a(r)	19	14.	a(r)	14	14.5	oo^2	33
15.	ur	15	15	oo^2	13	14.5	a(r)	33
16.	oi	12	16.	oi	9	16.	oi	21

Rank Difference Correlation: Clarke and Mt. Airy: Consonants .88, *PE* .02.
Vowels .92, *PE* .01.

TABLE 13

RELATIVE FREQUENCY OF CONSONANTS AND VOWELS (*a*) IN THE TEST MATERIALS; AND (*b*) IN DEWEY'S ANALYSIS OF 100,000 WORDS TAKEN FROM MODERN ENGLISH PROSE

The frequency of each sound is given in terms of per cent of total consonants, or per cent of total vowels.

	Test material this survey	100,000 words (Dewey)
t	9.6	11.8
n	9.0	12.0
m	7.9	4.6
th[0]	7.2	5.7
l	6.7	6.2
s	6.6	7.6
d	6.1	7.2
k	5.7	4.5
z	4.6	4.9
b	4.4	3.0
w	4.1	3.5
v	3.7	3.8
h	3.7	3.02
p	3.6	3.4
r	3.2	11.4
f	3.0	3.1
g	2.4	1.2
y	2.2	1.0
ng	1.5	1.6
sh	1.4	1.4
wh	.97*	—
ch	.92	.86
th[1]	.82	.61
j	.44	.73
zh	0	.08
–u–	22.8	14.0
–i–	14.8	19.2
–a–	7.4	9.6
ee	7.2	4.7
i-e	5.7	3.8
a–e	5.3	4.5
–e–	5.2	8.3
oo[1]	5.2	3.9
aw	4.6	3.1
o-e	4.2	4.0
a (r)	3.9	1.2
–o–	3.4	6.8
ou	2.5	1.4
ur	2.2*	—
oo[2]	1.7	1.8
oi	.86	.22

*Not included in Dewey's list. *Wh* counted as *h*; *ur* counted as *–u–* plus consonant final *r*.

Correlations Consonants: .89 *PE* .01

Vowels: .84 *PE* .12

according to its relative stability must be based upon the distribution of consonants and vowels in the test materials.

The phonemes in a random sample of 350 sentences were counted to determine the relative frequency of consonants and vowels in the test materials. The frequency of occurrence of each sound was converted into per cent of total consonants or vowels in the test materials. Rank lists of these percentages are presented in Table 13, ranked in the order of relative frequency. Similar lists taken from Dewey's material (3), representing an analysis of 100,000 words taken from representative types of modern English prose, are presented for comparison.[6] The two lists correspond fairly closely. There are some differences especially among the vowels. In the consonant lists the consonant *r* shows the widest variation among the phonemes of the two lists. Dewey's data show 11.4 per cent for *r* against 3.2 for *r* in the test materials used in this survey. This difference can be explained in part by the fact that no final *r*'s were included in our count because it is a general practice among teachers of the deaf to teach final *r* not as a consonant but as a vowel glide. The justification from a phonetic point of view for this practice is a debatable point and need not be considered here. Another reason for the slight differences between the two consonant counts is that no members of compound consonants were included in the count of the test materials. The reason for this has been discussed above. The larger differences between the vowel counts in the two lists are probably due to the relative simplicity of the test materials as compared with that used in the Dewey analysis. There is sufficient similarity, however, in both consonant and vowel lists to indicate that the test materials were not "loaded" with any particular types of sounds. The rank correlations between Dewey's lists and those of this survey are: consonant ranks, .89, *PE* .01; vowel ranks, .84, *PE* .12.

A ratio between the number of times an individual phoneme occurred in the test materials and the number of times this phoneme was malarticulated by the subjects provides an *Error Index* for each

---

[6]Dewey's percentages, showing each sound in terms of total consonants and vowels (3, p. 130, Table 17) have been converted so as to correspond to the data of the present survey. Each consonant is expressed in terms of per cent of total consonants and each vowel is expressed in terms of per cent of total vowels rather than taking both types of sounds together. The transcription used is that of the "Northampton Charts" (p. 304, Table *A*) (see also Yale, 24, p. 10) for convenience of teachers of the deaf.

TABLE 14

(1) Per cent of consonants and vowels in the test materials.
(2) Per cent of consonant and vowel errors.
(3) Error Index.
(4) Combined Error Index for the two school groups.

			Clarke			Mt. Airy			Combined	
		% cons. & vowels in tests	% errors	Error Index	*EI* rank	% errors	Error Index	*EI* rank	*EI* (sum)	*EI* rank
1	t	9.6	5.8	.60	20	5.1	.53	15	1.13	16
2	n	9.0	4.8	.53	21.5	1.4	.16	24	.69	23
3	m	7.9	3.5	.44	24	1.4	.18	23	.62	24
4	th[2]	7.2	5.5	.70	16.5	1.9	.26	19	.96	18
5	l	6.7	9.5	1.40	9.5	14.0	2.10	6	3.5	7
6	s	6.6	7.1	1.07	11	6.3	.95	12	2.02	12
7	d	6.1	9.1	1.50	8	18.4	3.10	1	4.6	2.5
8	k	5.7	4.0	.70	16.5	1.5	.25	20	.95	19.5
9	z	4.6	3.3	.75	15	6.4	1.40	8.5	2.15	11
10	b	4.4	7.0	1.60	6.5	11.7	2.70	3	4.3	4
11	w	4.1	2.5	.62	19	2.5	.62	14	1.24	14
12.5	v	3.7	2.4	.65	18	1.1	.30	18	.95	19.5
12.5	h	3.7	8.9	2.40	3	8.1	2.20	5	4.60	2.5
14	p	3.6	1.9	.53	21.5	.7	.19	22	.72	22
15	r	3.2	5.4	1.70	4.5	4.8	1.50	7	3.2	8
16	f	3.0	2.6	.87	13	.6	.20	21	1.07	17
17	g	2.4	3.9	1.60	6.5	5.6	2.40	4	4.0	5.5
18	y	2.2	3.2	1.40	9.5	2.8	1.30	10	2.7	10
19	–ng	1.5	.76	.50	23	.6	.40	16	.90	21
20	sh	1.4	3.9	2.80	2	1.7	1.2	11	4.0	5.5
21	wh	.97	.89	.92	12	.3	.31	17	1.23	15
22	ch	.92	1.6	1.70	4.5	1.3	1.40	8.5	3.10	9
23	th[1]	.82	.64	.78	14	.7	.85	13	1.63	13
24	j	.44	1.5	3.10	1	1.4	2.86	2	5.96	1
25	zh	0	0	0	—	0	0	—	0	—

TABLE 14 (*continued*)

		% cons. & vowels in tests	Clarke % errors	Clarke Error Index	Clarke *EI* rank	Mt. Airy % errors	Mt. Airy Error Index	Mt. Airy *EI* rank	Combined *EI* (sum)	Combined *EI* rank
1	-u-	22.8	5.1	.22	16	8.2	.36	16	.58	16
2	-i-	14.8	10.3	.70	14	21.6	1.46	5	2.16	11
3	-a-	7.4	6.1	.83	13	5.1	.69	14	1.52	14
4	ee	7.2	9.2	1.30	9	11.0	1.53	3	2.83	4
5	i-e	5.7	16.4	2.00	1	14.8	2.6	1	5.55	1
6	a-e	5.3	9.6	1.04	11	5.1	.96	8	2.00	12
7.5	-e-	5.2	6.9	1.30	8	7.8	1.5	4	2.80	5
7.5	oo^1	5.2	7.2	1.40	5.5	4.6	.89	10	2.29	9
9	aw	4.6	4.9	1.06	10	3.6	.78	12	1.84	13
10	o-e	4.2	6.8	1.60	3	3.0	.72	13	2.32	8
11	a(r)	3.9	2.6	.67	15	1.7	.44	15	1.11	15
12	-o-	3.4	4.7	1.38	7	4.2	1.24	7	2.62	6
13	ou	2.5	3.5	1.40	5.5	2.1	.84	11	2.24	10
14	ur	2.2	2.0	.91	12	4.4	2.00	2	2.91	3
15	oo^2	1.7	2.7	1.60	4	1.6	.94	9	2.54	7
16	oi	.86	1.6	1.86	2	1.1	1.28	6	3.14	2

Correlations: Clarke vs. Mt. Airy. Consonant *EI* .80, *PE* .05
Vowel *EI* .45, *PE* .14

Clarke vs. Combined. Consonant *EI* .91, *PE* .02
Vowel *EI* .71, *PE* .09

Mt. Airy vs. Combined. Consonant *EI* .95, *PE* .02
Vowel *EI* .85, *PE* .05

sound. This ratio, to be referred to as the *EI* ratio, can be used as a relative measure of the difficulty of each sound in the speech tests. If all phonemes had an *EI* ratio of 1, it would be an indication that all sounds were of equal difficulty, and that no one sound caused greater difficulty than any other. On the other hand, when sounds have *EI* ratios greater than 1, it is a clear indication that some factor other than mere frequency of occurrence is operative causing excessive errors on these particular sounds. When the *EI* ratio is less than 1 it means that this particular sound is giving less trouble than one would expect judging by its frequency of occurrence.

Table 14 shows the distribution of consonants and vowels in the test material; the per cent of error for each; and the *EI* ratio. The data for the two schools are given separately and then the *EI* ratios for the two groups are combined. There is a close agreement in consonant errors between the two schools as the rank correlation of .80, *PE* .05 indicates. There is less agreement for vowel errors; the rank correlation is .45, *PE* .14.

In the Error Index columns for the two schools, Table 14, there are 11 consonants each with *EI* ratios greater than 1. Furthermore, 10 of these 11 consonants are common to both columns; the fricatives *s* and *z* are interchanged. In other words, the same group of consonants which are more unstable in the speech of the Clarke pupils are also more difficult for Mt. Airy pupils. The group ot sounds which give trouble, i.e., those which are unstable in the speech of deaf children are: the sonant stops, *b, d,* and *g;* the affricatives *ch* and *j;* the fricatives *s, z, sh;* the liquids *l* and *r,* the semi-vowel *y,* and finally the aspirate *h.* The combined ranking of *EI* ratios for the two school groups (Table 14) gives them in the following order: *j, d, h, b, g, sh, l, r, ch, y, s,* and *z.* The five consonants which have the lowest *EI* ratios are *v, ng, p, n,* and *m.* Three of these five, *p, n,* and *m* appear among the lowest five in both columns. Each school group, therefore, experiences similar degrees of difficulty with identical consonants. The *EI* ratio, therefore, provides a reliable criterion for determining the relative difficulty experienced by deaf children in the accurate production of consonants, and for that reason, indicates the sounds upon which more careful practice and drill is needed.

There is less agreement on the vowels when ranked according to the Error Index. There are 7 vowels in the Mt. Airy rank list with *EI* ratios greater than 1, and 11 in the Clarke list. Of the 7

in the Mt. Airy list 3 are duplicated in the first 7 in the Clarke list. The combined ranking of *EI* vowel ratios, therefore, is less reliable than that of the consonants.

The significance of the *EI* ratio as a means of ranking consonants in their order of articulatory difficulty is clearly shown by the wide range of variation in magnitude of the ratios themselves and by the close agreement between the rankings of the same series of consonants by pupils in the two school groups (correlation .80). There are two possible explanations as to why one consonant is given a higher or lower ranking in the *EI* ratio than another: (*a*) The sound is more complicated, or more difficult from the standpoint of the articulatory coördinations involved in its production; or (*b*) all the sounds have the same degree of complexity but individual deaf pupils have become more skilled in the production and in the use of some consonants than in others. If the second of the above explanations were valid it would seem a rare coincidence to find certain consonants consistently among those more completely mastered while certain others were just as consistently malarticulated by 192 deaf children from two different school populations. If the second possibility were valid one would expect rather that no single consonant or group of consonants would be outstandingly difficult since the variations of individual pupils would obscure any trends in the analysis of the group as a whole. On the other hand, if the motor coördinations involved in the articulation of phonemes vary in degree of complexity this fact would be quickly manifested by the degree of accuracy with which all deaf children learn and use these sounds in their speech.[7]

2. *Teachers Ranking of Phonemes*

A second method of determining the relative difficulty of consonants and vowels was employed, based upon the composite judgment of experienced teachers of the deaf. Printed instructions with the consonants and vowels listed in a random order were given each teacher. The instructions were as follows:

---

[7]The thesis of Zipf (25), namely, that there is an inverse ratio, not necessarily proportional between the magnitude of complexity of phonemes and their frequency of occurrence in a language is partly supported by the data of this study. There is a negative correlation of —.48, *PE* .11 (Rank Method) between the consonants as ranked by the *EI* ratio, and as ranked according to frequency of occurrence in the speech material. When Dewey's phonetic counts (Table 13) are substituted for those of the present study the correlation is less, —.41, *PE* .12.

> Please rank the consonants and vowels listed below in the order of their *teaching difficulty*. Teaching difficulty means the relative difficulty deaf children face in learning these sounds, and the degree of stability of the sounds in the everyday speech of the deaf children. In other words, if you think some sounds are harder than others to teach, or some require more frequent correction as the children use them in their daily speech, rank them in that order. Please do not consult your colleagues in this matter.

Twenty teachers from the Primary, Intermediate, and Grammar Departments of Clarke School submitted ranking lists.

The individual rankings were combined into composite lists, one for each department, then a final composite list for the entire group. Teachers within the same department agree very well as to the order of difficulty of both consonants and vowels. The agreement between the three departments also was high. Rank correlation coefficients between the composite rankings of the three departments were as follows (Table *D*). Table 15 shows the composite teacher rankings of consonants and vowels.

TABLE *D*

	Consonants	Vowels
Primary vs. Intermediate	.91 *PE* .02	.82 *PE* .06
Primary vs. Grammar	.83 *PE* .05	.89 *PE* .03
Intermediate vs. Grammar	.83 *PE* .05	.86 *PE* .05

A comparison of the *EI* ratio ranking of phonemes with the teachers' ranking shows a rather low correlation. Table 15 shows the consonants and vowels ranked according to both methods. There is a correlation of .39, *PE* .12, between the consonant rankings. The agreement on vowels is slightly higher, correlation .47, *PE* .14. The coefficients of correlation between the *EI* rankings of the two schools taken separately and the teachers' composite ranking are listed at the bottom of Table 15.

The low correlation between the teachers' ranking of phonemes and the *EI* ranking indicate that, while teachers are able to agree among themselves as to the relative difficulty of consonants and vowels, their judgments do not agree with a scale of difficulty determined experimentally. The reliability of the *EI* ratio as a measure of relative difficulty of consonants is established both by the fact that there is a high agreement between *EI* rankings taken

TABLE 15

RANKINGS OF CONSONANTS AND VOWELS BY THE COMBINED ERROR INDEX RATIO (CLARKE AND MT. AIRY) AND A COMPOSITE OF 20 TEACHERS' RANKINGS

	Combined *EI* Ranks	Composite Teachers' Ranks
j	1	1
d	2.5	17
h	2.5	12
b	4	18
g	5.5	10
sh	5.5	8
l	7	4
r	8	5.5
ch	9	2
y	10	9
z	11	5.5
s	12	3
th[1]	13	22
w	14	19
wh	15	23
t	16	21
f	17	20
th[2]	18	14
k	19.5	11
v	19.5	15.5
ng	21	7
p	22	24
n	23	13
m	24	15.5
i-e	1	5
oi	2	8
ur	3	7
ee	4	3
–e–	5	1.5
–o–	6	10.5
oo[2]	7	10.5
o-e	8	9
oo[1]	9	15
ou	10	12
–i–	11	5
a-e	12	1.5
aw	13	13
–a–	14	6
a(r)	15	16
–u–	16	14

Correlations, Rank Diff. Method:	Consonants	Vowels
Clarke Teachers *vs.* Clarke *EI*	.47, *PE* .12	.03, *PE* —
Clarke Teachers vs. Mt. Airy *EI*	.44, *PE* .12	.61, *PE* .11
Clarke Teachers vs. Combined *EI*	.39, *PE* .12	.47, *PE* .14

from two distinct school groups and by the fact that it is based upon experimental data rather than judgments.

The *EI* ratios for vowels show considerably greater variability between the two schools than those of consonants. The correlation between Clarke and Mt. Airy for the two vowel rankings is .45, *PE* .14. There are two possible explanations for this variation in vowel rankings. In the first place, the articulatory movements for vowel positions in speech are less definite than those for consonants. For this reason the sensory cues from the vowel movements are more vague, and deaf children, who must depend upon the kinaesthetic cues for accuracy of movement, may make errors in articulating *any* vowel rather than particular ones. A second possibility is suggested by the fact that in analyzing the records the experimenters were less critical in assigning vowel errors than those of consonants. This would mean that many minor errors were not listed in the error categories.

The teachers' ranking of consonants and vowels agree with the *EI* rankings in certain isolated cases. For instance, all teachers agree that the consonants *j*, *ch*, *l*, *r*, and *s* are the most difficult sounds; they also agree that *p*, *f*, *m*, and *n* are relatively easy. The same teachers, however, place the sonant stops, *b*, *d*, and *g* near the bottom of their lists as being among the less difficult (Table 15). The *EI* ranking, on the other hand, places these three phonemes among the five most difficult consonants. Errors involving sonant stops are outstanding in the speech of all the more profoundly deaf pupils tested. It would appear, therefore, that teachers are either unaware of these errors in the speech of their children, or that they consider that the sound is learned when the child is able to give it as an individual element. There is also a wide discrepancy in the placement of the aspirate, *h*, by teachers and *EI* ratio ranking. Composite teacher ranking place *h* in 12th position while the *EI* ranking places it second. It is interesting to note the variation in the placement of the aspirate by teachers themselves: Primary teachers gave it 9th position, Intermediate teachers gave it a rank of 11, and Grammar Department teachers gave it a rank of 23. Primary teachers are more aware of the difficulties involved in teaching *h* than the others, although the older pupils very definitely fail to use the aspirate in approximately 9 per cent of attempts.[6]

[6]The aspirate *h* might have been given a separate category in the analysis of the data since it is totally different from all other consonants. It was included among consonants in this study in order to prevent further complicating the framework of error categories. The presence or absence

3. *Summary*

The method of analyzing articulatory errors makes it possible to determine the frequency of occurrence of errors on each of the consonants and vowels. By counting the actual number of phonemes in the test materials and converting individual consonant frequencies into percentages it is further possible to determine an Error Index which furnishes a means of ranking consonants and vowels on the basis of relative difficulty. The Error Index of a sound is defined as the ratio of the frequency of malarticulations of that sound and the frequency of its occurrence in the test materials. When the ratio is 1 the sound is of average difficulty; when it is greater than 1, the sound is of greater than average difficulty; when it is less than 1, the sound is of less than average difficulty.

A composite ranking of the same sounds by 20 teachers of the deaf, on the basis of the teaching difficulty of the sounds has a low correlation (.39) with the *EI* rankings. Teachers agree fairly well among themselves on the relative difficulty of consonants and vowels.

The *EI* ratios of vowels show a wider variation than those of consonants among the pupils of two schools for the deaf. The difference between the motor aspects of the two types of phonemes is suggested as the cause for this difference in variability.

---

of an aspirate preceding a word beginning with a vowel requires the speaker to make a distinction between two basic types of "vocal attack." Many profoundly deaf children either find this distinction extremely difficult or they have not been taught to distinguish between the two. Pupils who speak with breathy voices give the impression of beginning all vowels with aspirates and make neither physiological nor acoustic differences between such syllables as *harm* and *arm*, *ear* and *hear*. Auditors including classroom teachers take their cues from the context and supply the missing aspirate, or discount its presence when it is not implied by the context.

V. DISCUSSION OF RESULTS

A. Speech Intelligibility

Intelligibility must be the final criterion by which the speech of deaf children is to be judged. Articulatory errors were included in the final analysis of the data of this study only when at least one *audition error* could be attributed to them. A more rigid criterion would have yielded a far greater number of errors. There is some justification, however, for using intelligibility as a standard for determining whether or not variations from an objective "articulatory norm" should be counted as errors. A more rigid analysis would have ignored the very practical fact that often very poor speech was understood. The degree of toleration of the ear for varying degrees of phonetic inaccuracy, and the cues derived from sentence context are two factors which definitely limit the effect of articulatory errors upon intelligibility.

The intelligibility scores as determined by the methods described in previous pages probably are too low for both school groups. There are several factors other than the actual speech errors themselves which contributed to the low scores: (*a*) Losses due to the electrical transcription and reproduction; and (*b*) The absence of lip-reading cues in the auditing of the records. To these may be added the fact that unrelated sentences were used, thus limiting the contextual cues. These factors were counter-balanced to some extent by using as auditors persons familiar with the speech of the deaf and by allowing three repetitions of each sentence.

More accurate intelligibility averages can be approximated, however, by a bit of calculation. The loss in speech accuracy due to electrical transcription can be determined roughly by comparing intelligibility scores obtained by this method with those obtained in the annual speech tests (1, p. 35) taken the same year. When pupils read the test materials directly to the auditors, who are not watching the speaker, the average gain in intelligibility over the electrical transcription scores was 20 percentage points. This average gain was obtained from the same group of Clarke pupils used in this study.

It is also possible to check the effect of lip-reading cues upon speech intelligibility scores by means of the annual speech test data. In giving the tests four auditors are used, two of whom look at the speaker as well as listen; the other two merely listen. The "look-

listen" scores are usually higher than the "listen" scores by an average of 10-15 percentage points. This would indicate, therefore, that cues derived from lip-reading add considerably to the intelligibility scores.

Thus the sum of the loss due to the recording method used in this study and that due to the absence of lip-reading cues amounts to 30-35 percentage points. Adding this correction to the general averages of the Clarke pupils the average intelligibility level becomes something like 60-65 per cent, while that of Mt. Airy pupils is raised to 55-60 per cent. This estimate is probably nearer the true intelligibility of the pupils of both schools when their speech is considered from the standpoint of normal social intercourse.

B. Possible Causes of Speech Errors

The primary causes for the lack of intelligibility of the speech studied in this survey were the articulatory and rhythmic errors, the analysis of which has been presented. The seven consonant and five vowel error categories take into account practically all of the recurrent error types exhibited by the pupils tested. A close analysis of the several error categories reveals that they may be divided into two distinct groups on the basis of the motor processes in which the errors occur. It is of interest to examine the two groups of errors in terms of their frequency and relative importance for speech intelligibility. One group of error categories may be said to involve the inaccuracy, or the failure, of the articulatory processes themselves, while a second group involve a lack of coördination between articulatory processes and the breathing muscles. To the first group belong errors of consonant and vowel substitution, dropping components of compound consonants and diphthongs, nasality of consonants and vowels, and neutralization of vowels. To the second group belong errors involving the surd-sonant distinction, splitting compound consonants and diphthongs, non-functioning of releasing and arresting consonants, and finally, the malarticulation of abutting consonants. This second group of errors make up the larger proportion of errors in the speech of the pupils tested. Likewise, the categories of the second group are more important for speech intelligibility as determined by the criteria of relative importance. While not ignoring those errors which are purely articulatory, it may be said that the basic speech errors found in this survey are those

which represent failures of the speech mechanism as a whole, a lack of integration and coördination of the several component muscle groups *which make up the complex speech mechanism.* Errors of rhythm, also, the importance of which has been demonstrated, fall into this class.

The implications of these findings for the speech training program, and for methods of approach to this program are of signal importance. A very important question arises: Why do deaf children fail to achieve a higher degree of integration of the component elements of the speech mechanism? A logical answer to this question would be that the methods of speech training fail to take into account the complexities of the total speech mechanism and that sufficient practice is not afforded for the necessary integrative-coördinative processes. Speech *is mediated by a group of widely distributed organs coordinated* into a functional unit. Speech production is a secondary function of these organs each of which has had a primary function of its own in the vegetative and metabolic processes of the organism. Some of these processes, like respiration for instance, must be carried on concurrently with speech. Normal speakers learn quite unconsciously to make the necessary adjustments and modifications of the breathing mechanism early. Deaf children must be taught to make them, and to make them primarily in terms of movements without the acoustic cues to guide them. It is of vital importance, therefore, *that teachers have adequate knowledge of the detailed functioning of the several parts of the speech mechanism* in order to induce the proper coördinations necessary for fluent, intelligible speech. Emphasis upon one set of organs during the early training period, for instance the articulatory organs, and neglect of another set, the breathing muscles, will have far-reaching effects upon later speech developments. The integrative process for grouping syllables, accentuation, phrasing, should begin with the first speech exercises and be continued throughout the training period.

It is not possible to point to any definite cause, or set of causes, which may have produced the errors discussed in this paper. In lieu of stating definite causes, however, one may point out obvious factors which are operative in speech development of deaf children which may well lead to errors such as are herein discussed.

First of all, the absence of usable amounts of hearing, or impaired hearing, are admittedly primary factors in imperfect speech develop-

ment. *Speech errors and the degree of hearing loss are highly* correlated. The use of hearing aids for pupils who have usable amounts of *hearing definitely improves* the speech of those pupils who are given this advantage. That the factor of deafness is the sole cause for the relatively high frequency of speech errors, however, is open to question. The absence of hearing makes it necessary for the child to learn speech by artificial methods, and the accuracy of such speech will depend to a considerable degree upon the thoroughness of early speech training. The mental development of the child and the age at which speech instruction begins also affect speech development. Speech, once established, however, is not dependent upon hearing for its sensory control. Ample evidence for this is found in the fact that deafened children and adults can retain normal speech coördinations and intelligible speech after losing their hearing. The age of the person at the time of the incidence of deafness will determine in some *degree the amount of deterioration of speech following deafness.* Adults, in whom speech habits are definitely established, retain speech in a highly intelligible form. The speech of children, deafened at the ages of 3 to 6 years, will undergo serious deterioration. The stage of development of speech habits, therefore, is a primary determinant of the amount of deterioration to be expected following adventitious deafness. It is interesting to note, however, that speech deterioration following adventitious deafness is a matter of defective articulation and voice quality rather than that of rhythm and the basic coördinations between articulatory and breathing movements (9, pp. 22-24). Speech coördinations, therefore, can be controlled by kinaesthetic and tactile cues in the absence of hearing.

A second factor which must be considered in discussing the frequency *of errors in the speech of the pupils tested is the method* by which speech is taught. Speech teaching in schools for the deaf in America has followed rather closely the patterns set by the early pioneers in this field. The almost universally current method is known as the "Elements-" or the "Analytic Method." The basic assumption of this method is that spoken, like written, language consists of a discrete series of "events" or sounds strung together forming syllables, words, and phrases; and that these individual "elements" may be lifted out of their functional relationship to other sounds and taught separately as discrete entities. Once a sufficient number of elements has been learned the pupil will then be able to put them together and "build" words and phrases.

In the opinion of the writers the "Elements Method" of speech teaching is based upon a limited knowledge of the complexity of the speech mechanism as it functions in normal speech, and its assumptions are based upon the written rather than the spoken form of language.

In the first place speech is not a single series of events; it cannot be reduced to a single line of "sounds." There are two different and essential series of movements in speech, namely, the series of syllable pulsations of the chest-abdominal muscles, and the articulatory movements which produce the vowels and consonants (17, 18). These two series of movements occur concomitantly during the process of speaking and must be properly coördinated and integrated if normal speech is to be effected. The pulsations for the syllables are the dominant movement series and carry the rhythm of speech. Consonant and vowel movements, on the other hand, are the dependent series and are accessory to the syllable movements.

Syllables, rather than consonants and vowels are the basic *phonetic elements* in speech (15, pp. 27-32; 23, p. 85). Consonants are accessory movements which have definite function, (releasing, arresting, and linking of syllables) in speech. They have no independent existence and cannot be uttered properly apart from the syllables in which they function. Vowels may be uttered as single syllables but consonants never. The true consonantal function is that of opening and closing the oral canal for the release and arrest of syllables. The syllable pulsation is given acoustic qualities as the air activates the vocal folds thus producing the vocal tone which is modified by the oral cavities. The consonant movement may only partially close the oral cavity; or the vocal folds may be set into vibration during occlusion of the consonant; in either case the consonant itself takes on an acoustic quality. Indeed, the partially closed position of some consonants may be assumed and the resulting sound become a vowel: Thus the second *l* in *little,* or the *n* in *cotton*. The so-called semi-vowels *w* and *y* are extreme cases of this interchangeability of consonants and vowels. There is, however, a basic distinction between the consonantal and the vowel articulation in relation to the syllable. This distinction is based upon the nature of the articulatory movement (15, p. 39).

Properly articulated consonants involve accurate coördination between chest-abdominal muscles, which control the pulsations of the

syllables, and the lips, tongue, velum, and jaw movements. This coördination is vital for smooth, fluent speech. Consonants taught as separate "elements" often become, in the speech of deaf children, mere static positions, or slow transitory movements lacking the necessary dynamic qualities of articulate speech. When such "consonants" are combined with vowels to form syllables, words and phrases, the effect is that of mere juxtaposition of sounds strung together like beads on a string. The attention of the pupil is fixed upon the articulation of individual phonemes rather than upon the word or phrase of which the phonemes are inseparable parts (8, pp. 28-29).

Furthermore, the "Elements Method" of itself is not conducive to the development of normal speech rhythm. Stetson, (15, p. 205) in another connection, expressed the difficulty thus: "And it is not the case that one can first master the 'elements' of pronunciation, the 'sounds' and then set them in the rhythm." A series of sounds methodically strung together slows up the rate of syllable utterance and prevents the proper grouping of syllables into rhythmic patterns. Normally the rhythmic grouping of syllables within the phrase effects changes of duration, quality and intensity of consonants and vowels, and may even determine in which syllable the phonemes shall function. Since syllables, rather than individual sounds, carry the rhythm of speech (18, p. 246), the rhythmic grouping of syllables imposes definite limitations of duration and quality upon the individual phonemes. In the speech of many of the pupils analyzed in this study, however, the "chain of sounds" became the dominant series of events. Syllables, when apparent at all, recurred so slowly as to make rhythmic grouping impossible.

That the "analytic method" of speech teaching fails to produce intelligible speech in the case of some profoundly deaf pupils cannot be denied, and the low degree of intelligibility in many pupils is an indication that a careful reconsideration of methods of speech teaching is apparent. In this connection we may note the manner in which speech is acquired by normal hearing children. A long period of development precedes the fluent speech of normal hearing individuals. Imitation, through hearing, is of course the primary means of sensory control. Vowels and consonants as "elements" are ignored in the process. There is a considerable period during which the normal infant vocalizes and babbles in a completely automatic manner. In this process the speech mechanism is executing random movements

similar to those of the arms and legs. The child may be said to be "playing with his voice," or "experimenting" with sound. Finally words and phrases are spoken. These first attempts at speech are immature and faulty; the consonants and vowels are defective, but in later stages of development the errors are corrected. During the entire developmental period, however, the "experimental" efforts are carried on in terms of syllables, words and phrases (breath groups) rather than single phonemes. At no point in the development can it be said that the normal child analyzes or "breaks down" the phonetic unit, the syllable, into component parts.

The hearing child, during the speech development period, is exercising the entire speech mechanism. That is to say, the breathing muscles which control the column of air, the articulatory mechanism, including the laryngeal muscles, are brought into play simultaneously. Thus, the entire speech mechanism is undergoing a process of training; the several parts are becoming integrated and coördinated into a functional whole.

Deaf children, on the other hand, who are trained by the analytic method, receive training in articulation and elaborate exercises for tongue and lips while the breathing muscles during this period are either neglected or exercised separately. Thus, when the pupil has "mastered" a few consonants and vowels, he is expected to put them together to form syllables and words, and even phrases. But a very essential part of the speech mechanism has been left out of the preliminary training, and the resultant speech is halting, strained, and breathy.

There is sufficient evidence which indicates that congenitally deaf children pass through some of the early stages of babbling and automatic vocalizing. Often parents report that they were unaware the children were deaf until speech failed to develop at the usual age. Babbling and vocalizing may continue even through nursery school age. Sykes (20), who made observations on deaf children of nursery school age, reports that in their spontaneous vocalizations young deaf children used practically all of the vowel sounds and most of the consonants in standard English.

The early automatic vocalizations of deaf children, however, cannot in themselves lead to speech development since the absence of hearing prevents the imitative and corrective responses. These early efforts, however, can be supplemented, directed, and encouraged

both by parental care and by educational methods. The visual and tactile senses of the child may be utilized to encourage imitative speech responses, and the spontaneous vocalizations may be directed towards speech development. The revival in recent years of nursery schools for deaf children provides an excellent opportunity for experimentation with more natural methods of speech training. It is interesting to note that the formal "elements method" cannot be used with children of nursery school age, hence the informal natural (synthetic) method based upon imitation is indicated.

The so-called "Belgian Method" of speech teaching, widely used in Belgium and France and recently introduced in America is based upon the natural imitative approach. It is an attempt to carry over into speech teaching the methods long used in schools for the deaf for teaching reading (23, p. 97). It is an application of the psychological concept of "globalization" which was introduced into primary education by Decroly in Belgium in 1906 (23, p. 97). Herlin later introduced the method into the Belgian schools for the deaf (23, p. 99; 4, p. 8). The essential thing in this method is that the young deaf child is prepared for speech first by encouraging babbling and syllable exercises. Then he is encouraged to imitate words and phrases without first having learned the individual "elements" (4, 5, 21).

Haycock, a teacher of long experience in the English Schools for the deaf, advocates the "synthetic method" of speech teaching but warns (8, p. 27) that it is necessary to correct articulatory errors before they become fixed.

Obviously, deaf children must be taught correct articulatory responses if they are to learn to speak intelligibly. The fact that articulatory movements, however, are inseparably fused, from a functional standpoint, to the equally important movement which controls the column of air in speech cannot be overlooked, and must become a primary factor in speech teaching. A logical approach, therefore, to the problem of demutization of deaf children is one that makes the syllable the phonetic unit. Syllable movements which control articulatory movements and which release, arrest, and otherwise modify the air column must be taught simultaneously. The grouping of syllables into breath groups which means a controlled movement of expiration forms a vital part of such a program.

The findings of this investigation indicate a need for a general

reconsideration of methods of speech teaching in the light of modern experimental phonetics and speech development of normal children. Nursery schools for deaf children offer excellent experimental opportunities in this direction. There are encouraging indications that this movement is already in progress.

VI. SUMMARY AND CONCLUSIONS

Speech samples taken from 192 deaf pupils between the ages of 8 and 20 years from two oral schools for the deaf were analyzed to determine the degree of intelligibility, and to determine the frequency and relative importance of speech errors which interfered with intelligibility.

The purpose of the study was: (*a*) to identify speech errors, to classify recurrent errors into general types or categories, and to determine their frequencies; and (*b*) to determine the relative effects of each type of error upon speech intelligibility.

Each subject read 10 unrelated sentences which were recorded phonographically. The intelligibility of the sentences was obtained in quantitative terms by playing back the acoustic records to groups of auditors. The records were further analyzed to determine errors of articulation and rhythm.

Two general types of errors appeared: errors of articulation, involving both consonants and vowels, and errors of rhythm.

Consonant errors were classified into seven general types or categories as follows: (*a*) failure to distinguish between surd and sonant consonants; (*b*) consonant substitution; (*c*) excessive nasality; (*d*) malarticulation of compound consonants; (*e*) malarticulation of abutting consonants; (*f*) non-functioning of arresting consonants; and (*g*) non-functioning of releasing consonants.

Vowel errors were classified into five general categories: (*a*) vowel substitution; (*b*) malarticulation of diphthongs; (*c*) diphthongization of vowels; (*d*) neutralization of vowels; and (*e*) nasality of vowels.

Analysis of the speech samples from the standpoint of speech rhythm showed three types of rhythm: (*a*) sentences spoken with correct rhythm; (*b*) sentences spoken with abnormal rhythm; and (*c*) sentences spoken non-rhythmically.

The data were analyzed separately for the two school groups. The frequency and types of errors are presented to show their distribution according to degree of deafness and according to age.

Results show that approximately 21 per cent of all consonants and 12 per cent of all vowels were malarticulated by 192 deaf children.

Errors involving the surd-sonant distinction, compound consonants,

and the failure of the releasing consonants were found to be the most frequent of the error categories, and also the most important relative to speech intelligibility. The criteria of relative importance of consonant errors were: (*a*) degree of correlation between frequency of errors in each error category and speech intelligibility scores; and (*b*) the frequency of errors in terms of the number of errors possible in any category.

Vowel substitution and malarticulation of diphthongs were the most important of the vowel error types by the above criteria.

The frequency of errors of articulation and errors of rhythm increase with increased hearing-loss. A detailed analysis revealed that some of the individual error categories were more highly correlated with hearing-loss than others. Suggested explanations for these differences take into account: (*a*) acoustic and motor-phonetic differences in the structure of the phonemes involved; and (*b*) differences in the manner in which speech is learned by profoundly deaf and by partially deaf children.

Partially deaf pupils who have had consistent and systematic training with hearing aids make fewer errors than those with the same degree of hearing-loss who have not enjoyed this advantage.

Analysis of the data according to age shows that older pupils of the groups studied, especially the Clarke group, make a greater number of consonant errors and a lower average intelligibility score. Analysis of annual intelligibility test scores over a six-year period, however, indicate no general deterioration in speech with age. Individuals show losses or gains in these tests, but the general average over the six-year period is constant. A better explanation for the apparent deterioration with age which appears in this investigation is that the speech of older pupils of the Clarke group probably has been inferior from the beginning.

The importance of speech rhythm for speech intelligibility among deaf pupils is indicated by the fact that sentences spoken rhythmically correct have almost a 4 to 1 chance of being understood over those spoken with incorrect rhythm. Errors of rhythm, like errors of articulation, increase with increased hearing-loss. Sentences spoken *rhythmically* as distinguished from those spoken *arhythmically* tend to increase with age among pupils with all degrees of deafness. The *correct* form of English grouping and accentuation, however, does not appear to increase with age among the profoundly deaf pupils.

Comparison of the relative importance of speech rhythm, consonants, and vowels, for speech intelligibility reveals that the contributions of rhythm are just as great as those of consonants, and that both rhythm and consonants are of greater importance than vowels in determining speech intelligibility.

The methods of analysis of articulatory errors used in this study make it possible to compare one phoneme with another on the basis of relative difficulty and stability in the speech of 192 deaf pupils. The degree of difficulty of any consonant or vowel was determined by its *Error Index,* which is an expression of the ratio between number of times the sound was malarticulated and the frequency of occurrence of the sound in the speech test material. The Error Index ranking places each consonant and vowel in the order of its relative difficulty in the speech of deaf children. A second rank order, based on the combined judgments of 20 teachers of the deaf ranking each sound on the basis of its teaching difficulty, fails to agree with the Error Index ranking.

In conclusion, intelligibility must be the final criterion for judging the speech of the deaf. Intelligibility, or the lack of it, is determined largely by the frequency of errors of articulation and errors of rhythm. Those articulatory errors which most effectively interfere with intelligibility are those which result from a lack of coördination between the articulatory organs themselves and the muscles which control the column of air during speech. Speech rhythm, which is primarily a matter of grouping, accentuating and phrasing syllables, is also dependent upon a high degree of coördination between the several parts which make up the speech mechanism. Speech training methods, therefore, must take into account these facts. Methods which emphasize the training of the articulatory organs, to the neglect or exclusion of the breath-control mechanism, fail in the basic purpose of speech training; namely, that of establishing an integrated speech mechanism. Speech is a dynamic process; it cannot be broken down into static positions and isolated movements. An *analytic* method of speech teaching of itself, therefore, violates basic physiological and phonetic principles. A synthetic method in which the basic phonetic unit is the syllable is in keeping with these physiological and phonetic principles; it is the natural method by which hearing children acquire speech.

The findings of this study suggest the need for a reconsideration

of current methods of speech teaching in schools for the deaf. Hard-of-hearing pupils and those who have a usable amount of residual hearing should be taught by visual-auditory methods made possible by developments in modern hearing aids. Profoundly deaf pupils, to whom speech must remain a series of movements, must be given every opportunity for developing the complex movement coördinations which make up the speech process.

REFERENCES

1. 69th ANNUAL REPORT. The Clarke School for the Deaf. Northampton, Mass., 1936, 30-35.
2. BELL, A. G. The Mechanism of Speech. New York: Funk & Wagnalls, 1914. Pp. 133.
3. DEWEY, G. Relative Frequency of Speech Sounds. Cambridge, Harvard Univ. Press, 1923. Pp. 148.
4. DROUOT, M. E. Comment on Rend La Parole aux Sourds-Muets: La Méthode Belge de démutisation. Limoges, Imprimerie Nouvelle, 1930. Pp. 19.
5. ———, La préparation à la parole en méthode Belge. *Rev. Gén. d. l'Enseig. d. Sourds-Muets*, 1935, **36**, 155-158.
6. GUILDER, R. P., & HOPKINS, L. A. The importance of auditory functions studies in the educational program for the auditorily handicapped child. *Volta Rev.*, 1936, **38**, 69-74 and 149-155.
7. GUILFORD, J. P. Psychometric Methods. New York: McGraw-Hill, 1936. Pp. 566.
8. HAYCOCK, G. S. The Teaching of Speech. Stoke-on-Trent: Hill & Ainsworth, 1933. Pp. 302.
9. HUDGINS, C. V. A comparative study of the speech coördinations of deaf and normal subjects. *J. Genet Psychol.*, 1934, **44**, 1-48.
10. ———. Visual aids in the correction of speech. *Volta Rev.*, 1935 **37**, 637-643.
11. ———. & STETSON, R. H. Voicing of consonants by depression of the larynx. *Arch. Néerl. d. Phonét. Expér.*, 1935, **11**, 1-30.
12. RAWLINGS, C. G. A comparative study of the movements of the breathing muscles in speech and in quiet breathing of deaf and normal subjects. *Amer. Ann. Deaf*, 1935, **80**, 147-156; 1936, **81**, 136-150.
13. SCURI, D. Respirazione e Fonazione nei Sordomuti, *Rassegna di Educazione dei Sordomuti e Fonetica Biologica*, 1935, **14**, 82-113.
14. STETSON, R. H. A motor theory of rhythm and discrete succession. *Psychol. Rev.*, 1905, **12**, 250-350.
15. ———. Motor phonetics. *Arch. Néerl. d. Phonét. Expér.*, 1928, **3**, 1-216.
16. ———. Rhythm and rhyme. *Psychol. Rev. Monog.*, 1903, **4**, 413-466.
17. ———. Speech movements in action. *Transac. Amer. Laryngolog. Assoc.*, 1933, **55**, 29-41.
18. ———. The relation of the phoneme and the syllable. *Proc. 2nd. Internat. Cong. Phonet. Sci.* London: Cambridge Univ. Press, 1936. (Pp. 245-252).
19. STORY, A. J. The importance of the consonants in speech and speech-reading. *Volta Rev.*, 1905, **11**, 478-488.
20. SYKES, J. L. A study of the spontaneous vocalizations of young deaf children. (Studies in the Psychology of the Deaf, No. 1) *Psychol. Monog.*, 1940, **52**, 104-123.
21. THOLLON, B. L'Enseignement de la Parole aux Sourds-Muets: A Propos de la Méthode Belge. Paris: L'Institution Nationale des Sourds-Muets, 1928. Pp. 24.

22. VOELKER, C. H. A Preliminary Strobophotoscopic Study of the Speech of the Deaf. *Amer. Ann. Deaf*, 1935, 80, 243-259.

23. VIULLEMEY, P. La Préparation Psycho-Physiologique de la Parole chez L'Enfant Sourd. Nancy: Thomas, 1934. Pp. 275.

24. YALE, C. A. Formation and Development of Elementary English Sounds. Northampton: Metcalf, 1929. Pp. 43.

25. ZIPF, G. K. The Psycho-Biology of Language: An Introduction to Dynamic Philology. Boston: Houghton Mifflin, 1935. Pp. 329.

Preparation of Manuscripts for The Journal Press

A. DIVISIONS

I. THE TITLES OF JOURNAL ARTICLES AND THE MAJOR SUBDIVISIONS OF MONOGRAPHS ARE PRINTED IN TEN-POINT CAPS CENTERED ON THE PAGE

A. THE NEXT SUBDIVISION TITLE IS PRINTED IN CAPS AND SMALL CAPS CENTERED ON THE PAGE

1. *Then Italics, with Principal Words, Upper and Lower Case, Centered on the Page*

a. *Then italics, upper and lower case, 1-em run-in side head.*

(1). *Then italics, upper and lower case, 2-em run-in side head.*

(a). *Then italics, upper and lower case, 3-em run-in side head.*

[Further subdivision should be merged into the text without marginal indentation, and should be numbered with small letters.]

B. TABLES

Tables should always be typed on separate sheets of paper, numbered, and referred to in the text by number (Table 2). Tables should be so constructed that vertical lines are unnecessary.

C. CUTS

Drawings should be made on white drawing paper or on yellow or blue lined graph paper. Lettering should be large enough to be visible after photographic reduction. Lines should be heavy enough to be visible after photographic reduction. Drawings should be numbered, and referred to in the text by number (Figure 2). The legends should be typed on separate sheets of paper. In making drawings, only black India ink should be used. Never write instructions for the engraver, as your words may mean the opposite of what you intend.

D. REFERENCES

References should be segregated at the end of the paper, should be arranged in alphabetical order by author, numbered, and referred to in the text by number (2).

The proper form for a book reference is as follows:

2. DOE, J. The Preparation of Manuscripts. New York: Holt, 1938. Pp. 400.

The proper form for a journal reference is as follows:

2. DOE, J. The preparation of manuscripts. *J. Gen. Psychol.*, 1938, **19**, 550-575.

If in the text it is desirable to refer to a page, thus (2, p. 48).

E. GENERAL CAUTIONS

1. Use heavy typewriter paper, double-space the lines, and leave margins for editorial work.

2. Use as few footnotes as possible, but number such as you use, thus,[3]

[3]Footnote.

3. Retype any page on which written corrections have been made. Failure to do this will invite grief.

4. The Editor reads every manuscript carefully for meaning, and may modify punctuation in order to make meaning more clear. If you disagree with this on your page-proof, please initial your objection. The Editor will always yield on such matters to the author, but may not yield to the mechanical objections of a secretary.